

**DETECTION OF LARGE DELETIONS IN THE
FACTOR VIII (*F8*) AND VON WILLEBRAND
FACTOR (*VWF*) GENES USING MULTIPLEX
LIGATION-DEPENDENT PROBE AMPLIFICATION
(MLPA)**

A thesis submitted to the Manchester Metropolitan
University for the degree of Master of Philosophy in the
Faculty of Science and Engineering

March 2013

Michelle O'Brien

In collaboration with the Molecular Diagnostics Centre,
Department of Haematology, Central Manchester NHS
Foundation Trust

ABSTRACT

Haemophilia A and von Willebrand Disease (VWD) are disorders which can cause mild to severe bleeding in an individual. In many cases, knowledge of the molecular basis of these disorders is required to enable appropriate genetic counselling within families, carrier diagnosis and, in some cases, prenatal diagnosis. The current 'gold standard' for diagnosing patients with Haemophilia A and VWD disorders is by direct sequencing of genomic DNA to identify mutations. However this method may not reveal the presence of larger heterozygous deletions or duplications of *F8* and *VWF* due to masking by an unaffected allele. Therefore the genetic basis of disease in a proportion of cases may remain unresolved. Multiplex ligation-dependent probe amplification (MLPA) is a method based on gene dosage that can be applied to the detection of large gene deletions or duplications in Haemophilia and VWD, aiding effective genetic diagnosis.

This project developed and applied MLPA to investigate the molecular basis of disease in a selected cohort of patients diagnosed with Haemophilia A, type 1 and type 3 VWD, where standard direct sequencing protocols of *F8* and *VWF* had failed to reveal causative mutations, and to assess its application in the genetic diagnosis of these disorders. This extended previous work on the molecular basis of Haemophilia A and VWD carried out in the Molecular Diagnostics Centre (MDC) at Central Manchester NHS Foundation Trust.

MLPA confirmed the presence of large deletions in seven Haemophilia A index patients, two type 1 VWD patients, seven type 3 VWD patients and one possible duplication in a female haemophiliac. The presence of a masked gene abnormality was also confirmed in two carrier females.

This demonstrates that MLPA is a useful diagnostic tool for detecting or confirming larger scale mutations not suited to DNA sequence analysis. MLPA is an important addition to the range of molecular diagnostic tools for the investigation of these haemostatic disorders and may be utilised when conventional mutation screening has failed to identify a causative mutation.

ACKNOWLEDGEMENTS

I would like to thank the following people for their help and contribution to this project. Firstly I would like to thank my supervisors in the Molecular Diagnostics Centre at Manchester Royal Infirmary, Dr. Tony Cumming for his support and guidance throughout the project and Dr. Steve Keeney for his invaluable knowledge and advice concerning laboratory work and writing up this project. I would also like to thank my supervisor at Manchester Metropolitan University, Professor Bill Gilmore.

I would also like to thank all the staff in the Molecular Diagnostics Centre for all their help and support and for putting up with my constant procrastinating when it came to completing this project. I would especially like to thank Dr. Abida Awan for giving me time to complete this project and her constant enthusiasm for research and development and everything else in-between which always kept me going. I would like to thank Dr. Megan Sutherland for first proposing the idea of doing a project on bleeding disorders, I have enjoyed this project immensely, they will never be my first love but they have definitely grown on me. To my good friends (you know who you are), thanks for making it a pleasure to come to work everyday.

Thanks to my friends and family, who have always been there for me and provided me with support, encouragement and good times when I was feeling sorry for myself.

Not forgetting, my amazing husband James, who always knows how to cheer me up and make me laugh and supports me in a way no one else can. Thank you for everything, now we can finally relax on the sofa together and look forward to our new little arrival very soon!

CONTENTS

<u>Section</u>	<u>Page</u>
Abstract	ii
Acknowledgements	iii
Contents	iv
List of figures	xi
List of tables	xiv
Abbreviations and Symbols	xv
Amino Acid Abbreviations	xix
 Chapter 1: Introduction	
1.1 Normal Haemostasis	1
1.1.1 The Haemostatic response to Injury	1
1.1.2 Coagulation Pathway	2
1.1.2.1 Initiation	2
1.1.2.2 Amplification	3
1.1.2.3 Propagation	3
1.1.3 Regulation of Coagulation	4
1.1.4 Hypercoagulable States and Bleeding Disorders	5
1.2 Factor VIII	5
1.2.1 The role of Factor VIII in haemostasis	5
1.2.2 Physiology of FVIII	6
1.2.3 Molecular Biology of FVIII	6
1.3 von Willebrand Factor	8
1.3.1 VWF function in haemostasis and coagulation	8
1.3.2 Synthesis of VWF	9
1.3.3 Molecular Biology of VWF	10
1.4 Haemophilia A	12
1.4.1 Laboratory diagnosis of Haemophilia A	13
1.4.2 Treatment of Haemophilia A	14
1.4.3 Differential Diagnosis of Haemophilia A	15
1.4.3.1 Type 2N VWD	15
1.4.3.2 Combined Deficiency OF FV and FVIII	15

1.5 Molecular Pathology of Haemophilia A	16
1.5.1 Gross DNA Rearrangements	17
1.5.1.1 Common Partial Inversions of FVIII	17
1.5.1.2 Large Deletions	17
1.5.1.3 Insertions of Retrotransposons	18
1.5.1.4 Duplications	18
1.5.2 Point Mutations	19
1.5.2.1 Missense Mutations	19
1.5.2.2 Nonsense Mutations	20
1.5.2.3 mRNA Splice Site Mutations	20
1.5.2.4 Small Deletions/Insertions	20
1.6 Factor VIII Inhibitors	21
1.7 Inheritance of Haemophilia A	21
1.8 Genetic Diagnosis of Haemophilia A	22
1.8 Mosaicism	24
1.9 Detection of large scale deletions in heterozygous carriers	24
1.9.1. Multiplex Ligation-Dependent Probe Amplification (MLPA)	24
1.10 Molecular studies of large scale deletions/duplications in Haemophilia	25
1.11 von Willebrand Disease (VWD)	29
1.11.1 Clinical Manifestations	30
1.11.2 Bleeding history	30
1.12 Classification of VWD	30
1.12.1 Type 1 VWD	31
1.12.2 Type 2 VWD	32
1.12.2.1 Type 2A VWD	32
1.12.2.2 Type 2B VWD	32
1.12.2.3 Type 2M VWD	33
1.12.2.4 Type 2N VWD	33
1.12.3 Type 3 VWD	33
1.13 Laboratory Diagnosis of VWD	34
1.13.1 Screening Tests	34
1.13.1.1 Bleeding Time	34
1.13.1.2 Closure time	35

1.13.1.3 Platelet count	35
1.13.1.4 Activated Partial Thromboplastin Time (APTT)	35
1.13.2 Specific Tests	35
1.13.2.1 Factor VIII Activity (FVIII:C)	36
1.13.2.2 von Willebrand Factor Antigen (VWF:AG)	36
1.13.2.3 Ristocetin CoFactor Activity (VWF:RCo)	36
1.13.3 Discriminating Tests	37
1.13.3.1 Ristocetin-Induced Platelet Aggregation (RIPA) Assay	37
1.13.3.2 VWF Collagen Binding (VWF:CB)	37
1.13.3.3 VWF Multimer Analysis	38
1.13.3.4 VWF-Factor VIII Binding Assay	38
1.13.4 Treatment of VWD	38
1.14 Molecular Pathology of VWD	39
1.14.1 Type 1 VWD	39
1.14.2 Type 2 VWD	41
1.14.2.1 Type 2A	41
1.14.2.2 Type 2B VWD	41
1.14.2.3 Type 2M VWD	41
1.14.2.4 Type 2N VWD	42
1.14.3 Type 3 VWD	43
1.15 Genetic Testing in VWD	46
1.15.1 Diagnostic Strategy	46
1.16 Detection of large scale heterozygous deletions	47
1.17 Molecular studies of large deletion in VWD using MLPA	48
1.18 Project Aims and Objectives	50
1.19 Previous studies related to this project	51
1.19.1 General Mutation Detection Strategies for Haemophilia A	51
1.19.1.1 Intron 22 inversion detection by Long PCR	51
1.19.1.2 PCR Amplification of the <i>F8</i> Intron 1 Inversion	51
1.19.1.3 Direct DNA Sequencing	52
1.19.1.4 Detection of Partial Gene Deletions	52
1.19.2 General Mutation Detection Strategies for VWD	53
1.19.2.1 Type 1 VWD	53
1.19.2.2 Type 3 VWD	54

1.19.2.3 Screening of <i>VWF</i>	54
1.19.2.4 Identification of the <i>VWF</i> exon 4 and 5 deletion	55
Chapter 2: Materials and Methods	
1.2 Selection of Patients	56
2.1.1 Haemophilia A	56
2.1.2 Type 1 VWD	56
2.1.3 Type 3 VWD	57
2.2 Phenotypic Data	57
2.3 Sample Collection and Processing	58
2.4 MLPA® Technology	58
2.4.1. MLPA Probe Chemistry	59
2.4.2 MLPA Reaction	60
2.4.3 Experimental set up	61
2.4.3.1 Reference samples	61
2.4.3.2 Sample treatment	61
2.5 SALSA® MLPA® Kits	62
2.6 MLPA Reaction method	63
2.6.1. DNA Denaturation (Day 1)	64
2.8.1.1 RNase treatment	64
2.6.2. Hybridisation reaction (Day 1)	64
2.6.3. Ligation reaction (Day 2)	65
2.6.4. PCR reaction (Day 2)	66
2.6.4.1 One tube protocol	66
2.6.4.2 Two-tube protocol	67
2.6.5 Fragment Separation by Capillary Electrophoresis (CE)	68
2.6.6 Overview of fragment analysis	69
2.7 Analysis of Data	70
2.7.1 Evaluating data	70
2.8 Interpretation of Results	72
Chapter 3: Results	
3.1 Choice of Method	77
3.2 MLPA Experiment Optimisation	77

3.2.1 Results of the Evaporation test	77
3.2.2 Evaluation of the MRC-Holland MLPA method using control samples	78
3.2.3 Assessment of sample quality	80
3.2.4 Optimisation of the MLPA thermal cyclers programme	80
3.2.5 Internal quality controls and Non-specific peaks	83
3.2.6 Bioinformatic analysis	85
3.3. Results of MLPA analysis in the Haemophilia A patient group	86
3.3.1 Mild haemophilia A patients with no causative mutation identified	88
3.3.2 Severe haemophilia A patients with large deletions and carrier relatives	91
3.3.2.1 Patients 5 and 6	91
3.3.2.2 Patients 7 and 8	93
3.3.2.3 Patients 9 and 10	95
3.3.2.4 Patient 11	97
3.3.2.5 Patient 12	98
3.3.2.6 Patient 13	99
3.3.2.7 Patient 14	100
3.3.2.8 Patient 15	101
3.3.2.9 Patients 16 and 17	102
3.3.2.10 Patients 18 and 19	105
3.4 Results of MLPA analysis in the Type 3 VWD patient group	107
3.4.1 Investigation of Type 3 VWD patients with exon 4-5 deletion by MLPA	109
3.4.2 Confirmation of homozygosity for exon 4-5 deletion by MLPA	110
3.4.3 Confirmation of heterozygosity for exon 4-5 deletion by MLPA	112
3.4.4 Investigation of type 3 patients with no causative mutation by MLPA	115
3.5 Results of MLPA analysis in the Type 1 VWD patient group	118
3.5.1 Investigation of type 1 VWD index patients with exon 4-5 deletion by MLPA	121
3.5.2 Investigation of type 1 VWD index patients with no causative mutation by MLPA	122

Chapter 4: Discussion

4.1 Introduction	125
4.2 MLPA method evaluation	126
4.2.1 Sample Treatment	128
4.2.2 Method Optimisation	129
4.2.3 Data analysis	130
4.2.4 Result interpretation	131
4.2.5 Advantages and Limitations of MLPA	132
4.3 Haemophilia A patients	134
4.3.1 Severe Haemophilia A patients	135
4.3.1.1 Application of direct assays	135
4.3.1.2 Determination of carrier status	136
4.3.1.3 Identification of large duplications	137
4.3.1.4 Presence of large deletions and link with development of inhibitors	138
4.3.2 Mild Haemophilia A patients	139
4.4 VWD patients	142
4.4.1 Confirmation of the Exon 4-5 deletion in VWD patients	143
4.4.2 Confirmation of heterozygosity for the exon 4-5 deletion	143
4.4.3 Confirmation of the exon 4-5 deletion in type 1 VWD patients	143
4.4.4 Overview of the exon 4-5 deletion	144
4.4.5 Type 1 VWD with no causative mutation	145
4.5 Mutation Detection Strategies – The role of MLPA	148
4.5.1 Proposed Mutation Screening Algorithm incorporating MLPA for Haemophilia A and VWD	148
4.5.1.1 Haemophilia A	148
4.5.1.2 VWD	150
4.6 Conclusions and future work	151

References	155
-------------------	-----

Appendices

Appendix 1: SALSA MLPA P178-B1 F8 Probemix	173
Appendix 2a: SALSA MLPA P011-B1 (Kit 1) <i>VWF</i> Probemix	175
Appendix 2b: SALSA MLPA P012-B1 (Kit 2) <i>VWF</i> Probemix	176

Appendix 3: <i>VWF</i> sequence with probe locations	177
Appendix 4: <i>F8</i> sequence with probe locations	190

LIST OF FIGURES

<u>Number</u>	<u>Figure Title</u>	<u>Page</u>
Figure 1.1	The three stage cell-based model of coagulation	4
Figure 1.2	The lifespan of FVIII	6
Figure 1.3a	Schematic representation of the chromosomal localisation of <i>F8</i>	7
Figure 1.3b	Exonic structure of <i>F8</i>	7
Figure 1.3c	Protein domains of FVIII	7
Figure 1.4	Mechanism of Platelet Adhesion to VWF	9
Figure 1.5	Schematic representation of the functional domains and binding sites of the VWF Protein	12
Figure 1.6	Inheritance in haemophilia	23
Figure 1.7	Diagram representing the distribution of type 1 VWD mutations in <i>VWF</i>	40
Figure 1.8	Diagram representing the distribution of type 2 VWD mutations in <i>VWF</i>	42
Figure 1.9	The distribution of type 3 VWD mutations in <i>VWF</i>	44
Figure 2.1	MLPA® probe set	59
Figure 2.2	MLPA reaction overview	60
Figure 2.3	Raw peak patterns visualised using GeneMapper Software	72
Figure 2.4	Example of the NGRl spreadsheet	74
Figure 2.5	Example of a histogram from the P178-B1 MLPA spreadsheet summarising the mean DQ data	75
Figure 3.1a	Initial test control sample histograms for MLPA kits – <i>F8</i>	79
Figure 3.1b	Initial test control sample histograms for MLPA kits – <i>VWF</i> kit 1	79
Figure 3.1c	Initial test control sample histograms for MLPA kits – <i>VWF</i> kit 2	79
Figure 3.2	Example of a poor quality sample used for MLPA analysis of <i>F8</i>	80
Figure 3.3	Genotype plots for a sample tested in parallel to compare the MRC-Holland and Edinburgh parameters	83

Figure 3.4	Example of a non-specific peak observed in the 64-82 nt range.	84
Figure 3.5	Examples of a non-specific peak at ~87 nt observed in two different patient samples	85
Figure 3.6a	MLPA Histogram – Patient 1	89
Figure 3.6b	MLPA Histogram – Patient 2	90
Figure 3.6c	MLPA Histogram – Patient 3	90
Figure 3.6d	MLPA Histogram – Patient 4	90
Figure 3.7a	Image of the agarose gel where exons 16-21 failed to amplify following standard PCR amplification	92
Figure 3.7b	Sequence trace for the partial gene deletion (g.116547_123812del7356) showing the exon 15 breakpoint	92
Figure 3.8a	MLPA Histogram – Patient 5	93
Figure 3.8b	MLPA Histogram – Patient 6	93
Figure 3.9a	MLPA Histogram – Patient 7	94
Figure 3.9b	MLPA Histogram – Patient 8	95
Figure 3.10a	MLPA Histogram – Patient 9	96
Figure 3.10b	MLPA Histogram – Patient 10	96
Figure 3.11	MLPA Histogram – Patient 11	97
Figure 3.12	MLPA Histogram – Patient 12	98
Figure 3.13	MLPA Histogram – Patient 13	100
Figure 3.14	MLPA Histogram – Patient 14	101
Figure 3.15	MLPA Histogram – Patient 15	102
Figure 3.16	Family tree for Patients 16 and 17	103
Figure 3.17a	MLPA Histogram – Patient 16	104
Figure 3.17b	MLPA Histogram – Patient 17	104
Figure 3.18a	MLPA Histogram – Patient 18	106
Figure 3.18b	MLPA Histogram – Patient 19	106
Figure 3.19	<i>F8</i> transcript for exon 17	107
Figure 3.20	Example of the identical exon 4-5 deletion breakpoints in patients homozygous and heterozygous for the mutation	110
Figure 3.21a	MLPA Histogram – Patient T3-5 (Kit 1)	111
Figure 3.21b	MLPA Histogram – Patient T3-5 (Kit 2)	112
Figure 3.22a	MLPA Histogram – Patient T3-1 (Kit 1)	115

Figure 3.22b	MLPA Histogram – Patient T3-1 (Kit 2)	115
Figure 3.23	MLPA Histogram – Patient T3-9 (Kit 2)	118
Figure 3.24	<i>VWF</i> transcript for exon 30	118
Figure 3.25a	MLPA Histogram – Patient T1-3 (Kit 1)	122
Figure 3.25b	MLPA Histogram – Patient T1-3 (Kit 2)	122
Figure 3.26a	MLPA Histogram – Patient T1-16 (Kit 1)	124
Figure 3.26b	MLPA Histogram – Patient T1-16 (Kit 2)	124
Figure 4.1	Mutation screening algorithm incorporating MLPA for haemophilia A patients	149
Figure 4.2	Mutation screening algorithm incorporating MLPA for VWD patients	151

LIST OF TABLES

<u>Number</u>	<u>Table Title</u>	<u>Page</u>
Table 1.1	Classification of VWD	34
Table 2.1	Components of SALSA® MLPA® kit	63
Table 2.2	Thermocycler Program for the MLPA Denaturation, Hybridisation and Ligation Reactions	66
Table 2.3	Thermocycler Program for the MLPA PCR Reaction	67
Table 2.4	Internal Quality Control Fragments	70
Table 2.5	Typical ranges for Dosage Quotients	75
Table 3.1	Results of the evaporation test comparing two brands of tubes	78
Table 3.2	Thermocycler Program for the MLPA Denaturation, Hybridisation, Ligation and PCR Reactions as per Edinburgh laboratory	81
Table 3.3	Haemophilia A patients and query carrier relatives investigated by MLPA analysis following routine <i>F8</i> screening	87
Table 3.4	Type 3 VWD patients investigated by MLPA analysis following screening for mutations within the <i>VWF</i>	108
Table 3.5	Type 1 VWD index patients investigated by MLPA analysis following screening for mutations within <i>VWF</i>	120

LIST OF ABBREVIATIONS AND SYMBOLS

<u>Abbreviation</u>	<u>Description</u>
aa	Amino acid
ABI	Applied Biosystems
ADAMTS13	A disintegrin and metalloprotease with thrombospondin repeat
ADP	Adenosine diphosphate
APTT	Activated partial thromboplastin time
AT	Antithrombin
bp	Base pairs
BSS	Bleeding score severity
BT	Bleeding time
c.	Nomenclature for cDNA level
°C	Degrees Celsius
CCD	Charge coupled device
cDNA	Complementary DNA
CMFT	Central Manchester Foundation Trust
CNV	Copy number variation
CPA	Clinical Pathology Accreditation
DDAVP	1-Desamino-8-D-Arginine Vasopressin
Delmap	Deletion mapping primer
dH ₂ O	Distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleoside triphosphates
DQ	Dosage Quotient
EC	Endothelial cell
EDTA	Ethylenediaminetetraacetic acid
Ex4-5del	Deletion of <i>VWF</i> exons 4 and 5
FII	Coagulation factor II
FIX	Coagulation factor IX
FIXa	Activated factor IX
FVa	Activated factor V

FVIIa	Activated factor VII
FVIII	Coagulation factor VIII
FVIIIa	Activated factor VIII
FVIII:C	FVIII activity
FXa	Activated factor X
FXa/FVa	Factor X / factor V complex
FXI	Coagulation factor XI
FXIa	Activated factor XI
FXII	Coagulation factor XII
FXIII	Coagulation factor XIII
FXIIIa	Activated factor XIII
F5F8D	Combine deficiency of factor V and factor VIII
<i>F8</i>	FVIII gene
<i>F9</i>	FIX gene
FISH	Fluorescent in situ hybridisation
Gp	Glycoprotein
GpIb	Glycoprotein Ib
GpIIb/IIIa	Glycoprotein IIb/IIIa complex
GP6	Glycoprotein VI
HAMSTeRS	Haemophilia A Mutation Database and FVIII Resource site
HGVS	Human genome variation society
HMW	High molecular weight
HMWK	High molecular weight kininogen
ISTH	International society for thrombosis and haemostasis
ISTH SSC	ISTH scientific sub-committee
int1h	intron 1 homologous
IU/dL	International units per decilitre
IU/mL	International units per millilitre
kb	Kilobases
kDa	Kilo Daltons
LA PCR	Long and accurate PCR
LC	Liquid chromatography

LINEs	Long interspersed repeat elements
<i>LMAN-1</i>	Lectin Mannose Binding Protein gene
LREC	Local Research Ethics Committee
µg	Micrograms
µL	Microlitres
µM	Micromolar
M	Molar
<i>MCFD2</i>	Multiple Coagulation Factor Deficiency 2 gene
MDC	Molecular Diagnostics Centre
Min	Minutes
mg/mL	Milligrams per millilitre
mL	Millilitres
MgCl ₂	Magnesium chloride
MLPA	Multiplex Ligation-dependent Probe Amplification
mRNA	Messenger RNA
MREC	Multi-Centre Research Ethic Committee
MRI	Manchester Royal Infirmary
N13F	N13 forward primer
N13R	N13 reverse primer
NEQAS	National External Quality Assurance Scheme
NGRL	National Genetics Reference Laboratory
NHS	National Health Service
NW	North West
ng	Nanograms
nt	Nucleotides
p.	Nomenclature for protein level
PCR	Polymerase chain reaction
PFA-100™	Platelet function analyser
PT	Prothrombin time
RNA	Ribonucleic acid
RGD	Arginine – Glycine – Aspartic Acid
RIPA	Ristocetin induced platelet aggregation
RT	Room temperature
s	seconds

SINEs	Short interspersed repeat elements
SNP	Single nucleotide polymorphism
<i>Taq</i>	<i>Thermus aquaticus</i>
TBE	Tris borate EDTA buffer
TCT	Thrombin clotting time
TE	Tris EDTA
TF	Tissue factor
TFPI	Tissue factor pathway inhibitor
TF/FVIIa	Tissue factor / factor VII complex
TT	Thrombin time
TFPI	Tissue factor pathway inhibitor
UKHCDO	United Kingdom Haemophilia Centre Doctor's Organisation
UK NEQAS	United Kingdom National External Quality Assurance Scheme
UTR	Untranslated region
VWD	von Willebrand disease
VWF	von Willebrand factor
<i>VWF</i>	von Willebrand factor gene
<i>VWFP</i>	von Willebrand factor pseudogene
VWF-FVIII	VWF-FVIII complex
VWF:Ag	von Willebrand factor antigen
VWF:CB	von Willebrand factor collagen binding
VWF:FVIII B	von Willebrand factor factor VIII binding
VWF:RCo	von Willebrand factor ristocetin co-factor

AMINO ACID ABBREVIATIONS

<u>Amino acid</u>	<u>Single letter code</u>	<u>Three letter code</u>
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic acid	D	Asp
Cysteine	C	Cys
Glutamic acid	E	Glu
Glutamine	Q	Gln
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val
Stop codon	*	

1.1 NORMAL HAEMOSTASIS

Normal haemostasis is responsible for ensuring the free flow of blood within the circulatory system. The efficient flow of blood and prevention and control of bleeding from damaged blood vessels is maintained by haemostatic mechanisms which depend on a series of complex interactions involving the vascular endothelium, circulating platelets and activation of coagulation pathways. The balance of normal haemostasis maintains the equilibrium between keeping the blood fluid and preventing excessive activation of the procoagulatory components leading to intravascular thrombosis.

1.1.1 The Haemostatic Response to Injury

The initial haemostatic response to vascular damage is an immediate vasoconstriction of the injured vessel and constriction of adjacent small arteries and arterioles to reduce the flow of blood to the area of injury. The reduced blood flow allows contact activation of platelets and coagulation factors. The lumen of blood vessels is lined with endothelial cells (ECs) which form an impermeable layer preventing passive transfer of blood cells, except in the case of inflammation. Within the sub endothelial layer are a number of components which have specific functions in both primary and secondary haemostasis.

Following damage to the endothelial layer, the sub endothelial layer is exposed. This leads to activated platelets which adhere to exposed subendothelial fibrils, potentiated by von Willebrand factor (VWF). Collagen fibers are the most thrombogenic components of the extracellular matrix. Glycoprotein VI (GpVI) expressed on the surface of platelets binds to exposed collagen inducing platelet adhesion, degranulation and aggregation. The platelets release chemical messengers such as adenosine diphosphate (ADP), serotonin and thromboxane A₂ which cause them to aggregate to form a platelet plug covering the exposed connective tissue and providing temporary control of bleeding.

1.1.2 Coagulation Pathway

The complete response is achieved by the process of secondary haemostasis. This involves the activation of the coagulation system leading to the generation of fibrin strands which are laid down between platelets and reinforce the platelet plug. The classical model of blood coagulation was represented by a coagulation cascade mediated via two pathways - the contact activation pathway (intrinsic) and the tissue factor pathway (extrinsic) which were initiated by separate mechanisms and converged on a common pathway leading to clot formation (Davie and Ratnoff 1964; Macfarlane 1964). The intrinsic system explained the reaction of enzymes intravascular to the blood, involving Factor XII, high-molecular-weight kininogen (HMWK), prekallikrein and Factor XI (FXI), Factor IX (FIX) and Factor VIII (FVIII) whilst the extrinsic system was named because it was activated by factors external to the blood system, i.e. tissue factor (TF). The pathways merged with the binding of Factor X (FXa) to Factor V (FVa) to form prothrombinase leading to the subsequent conversion of prothrombin into thrombin and the thrombin mediated formation of fibrin from fibrinogen. The fibrin clot was subsequently stabilised by the action of the activated Factor XIII (FXIIIa).

More recently, a cell based model has been described which explains the pathophysiology of the mechanisms leading to clot formation, highlighting the roles of different cells and phospholipid surfaces for the localisation of the coagulation network to the site of vascular injury. According to the cell based model, coagulation occurs in three different overlapping stages of initiation, amplification and propagation on tissue factor bearing cells and platelets (Hoffman 2003) (see Figure 1.1).

1.1.2.1 Initiation

In response to vascular damage, the coagulation process is initiated by the release of tissue factor (TF) by the endothelial cells. The circulating zymogen Factor VII (FVII) is exposed to cells expressing TF on their surface and is activated to become the protease Factor VIIa (FVIIa). FVIIa forms a high affinity complex with TF (TF/FVIIa) which leads to the activation of Factors IX (FIXa) and X (FXa). FXa then activate FV and forms a prothrombinase

complex (FXa/FVa) which generates trace amounts of thrombin from prothrombin.

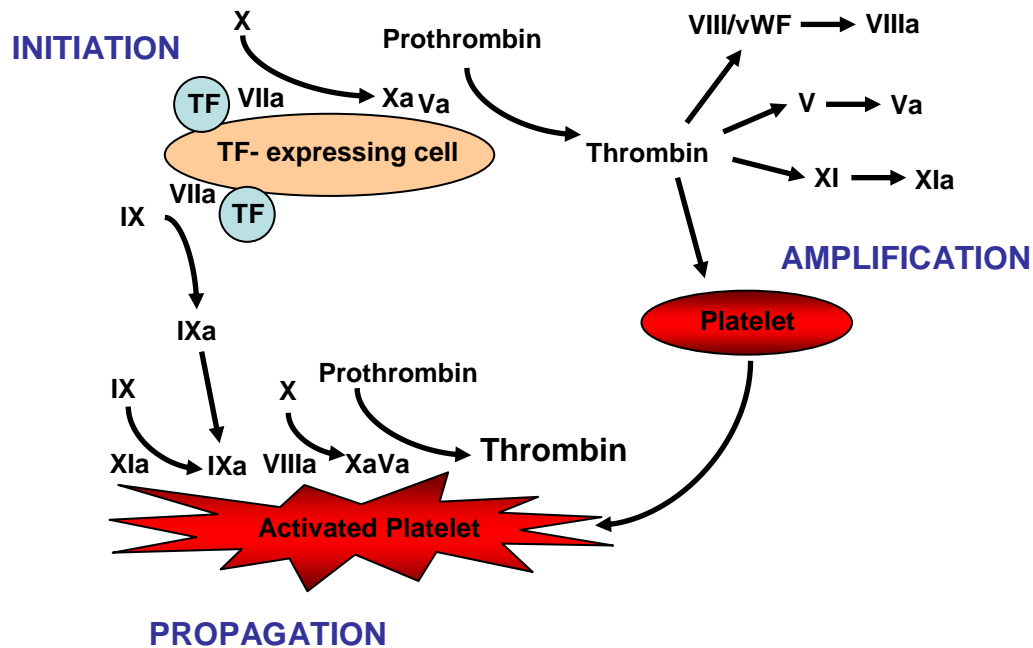
1.1.2.2 Amplification

Under normal conditions, haemostatic components in the vascular tissue are not able to leave the endothelial layer due to their size. However, upon vessel wall injury, platelets, FVIII and VWF are able to leave the vascular tissue and come into close contact with thrombin generated in the initiation stage. Platelets become activated by the small amount of thrombin leading to exposure of binding sites and receptors on their surfaces for activated clotting factors. The activated platelets release increased quantities of FV which becomes activated by thrombin. Thrombin also cleaves FVIII from its stabiliser molecule VWF activating it on the platelet surface and activates Factor XI (FXIa) which binds to the surface of activated platelets.

1.1.2.3 Propagation

The propagation stage occurs on the phospholipid surface of activated platelets that have joined the site of vascular injury. FXIa activates increased amounts of Factor IX (FIXa) which forms a complex with its cofactor FVIIIa on the surface of platelets known as the 'tenase' complex. This complex activates sufficient FXa which in turn forms the 'prothrombinase' complex with FVa resulting in the explosive generation of thrombin from prothrombin. This ultimately leads to generation of a fibrin clot by cleaving fibrinogen to fibrin monomers which polymerise to form a platelet plug. The clot is stabilised by Factor XIII (FXIIIa) which is also activated by the large burst of thrombin.

Figure 1.1: The three stage cell-based model of coagulation (figure adapted from Jurlander, Thim et al. 2001). The initiation phase results in an initial burst of thrombin generation followed by amplification of the procoagulant signal, which leads to propagation of the procoagulant response causing increased thrombin generation and the formation of a stable fibrin plug.



1.1.3 Regulation of Coagulation

Normal haemostasis is maintained by surface expression of molecules on endothelial cells, circulating inhibitors of plasma proteins and negative feedback mechanisms. Tissue factor pathway inhibitor (TFPI) and antithrombin (AT) are both key regulators of coagulation. TFPI is regulated by binding to glycosaminoglycans on the surfaces of ECs while AT is bound to heparin sulphate proteoglycans on the surface of ECs. Together they are able to inhibit any FXa that dissociates from the TF-bearing cell (on which it should form the prothrombinase complex with FVa).

Thrombin plays a key role in the coagulation system because as well as its procoagulant role, it also has anticoagulant properties since it activates factors involved in fibrinolysis to dissolve the fibrin clot. Thrombin dispersing from the site of vascular injury forms a complex with thrombomodulin (expressed on

ECs) to activate protein C which when bound to protein S inactivates FVIIIa and FVa, thus restricting clotting factors to the site of injury.

1.1.4 Hypercoagulable States and Bleeding Disorders

Abnormal bleeding is the result of defective vascular function or disturbed haemostasis after injury. The great variety of congenital bleeding disorders provides clear evidence of the key role played by single components of normal haemostasis. The bleeding tendency in patients with disorders such as von Willebrand disease (VWD) and the haemophilias demonstrates that all of the components are required for normal haemostasis, as a severe deficiency of one cannot be compensated for by normal activity/concentration of the others (Rasche 2001).

1.2 FACTOR VIII

FVIII is a complex plasma glycoprotein of 2351 amino acids that has an essential role in the haemostatic system. It is involved in the intrinsic pathway of the coagulation cascade where it functions as a cofactor for the factor X-activating complex, which consists of FVIIIa, the serine protease FIXa, calcium ions and phospholipid membranes (Mann 1999).

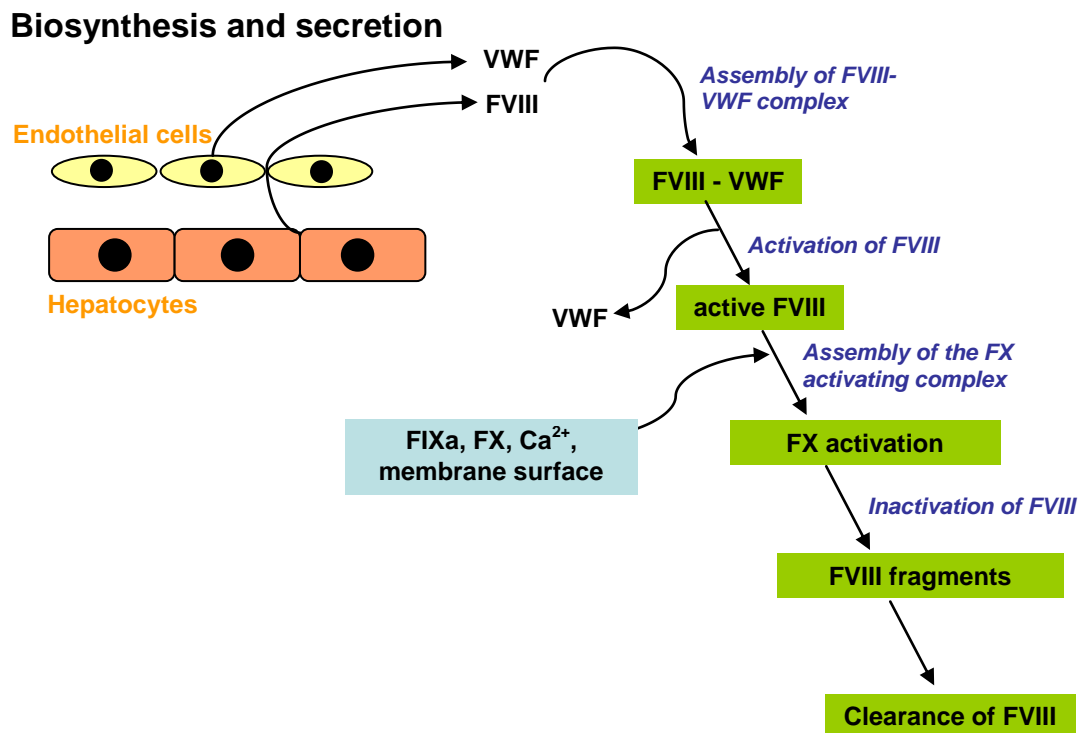
1.2.1 The role of Factor VIII in haemostasis

The bleeding associated with haemophilia patients is due to failure of the secondary phase of haemostasis. The formation of the platelet plug occurs normally but there is defective stabilisation by fibrin due to inadequate thrombin generation. This is because FVIII and FIX are now known to be central to the blood coagulation process and for generation of adequate amounts of thrombin. Following blood vessel injury, TF activates FVII which mediates generation of FXa, and FVIII and FIX are required for amplification of this production to enable completion of the coagulation process. If either FVIII or FIX are absent, bleeding will arise due to the insufficient amplification and generation of FXa to maintain haemostasis (Bolton-Maggs and Pasi 2003).

1.2.2 Physiology of FVIII

The major source for plasma FVIII is the hepatocytes although the kidney, sinusoidal endothelial cells and lymphatic tissues can also synthesise small amounts of FVIII (Hollestelle, Thinnes et al. 2001). It is produced as an inactive single-chain protein but following endoproteolytic processing, it is released into the circulation as a set of heterodimeric proteins. FVIII freely interacts with VWF, circulating in a non-covalent complex which protects it from premature proteolytic degradation and concentrates it at sites of vascular injury (Lenting, van Mourik et al. 1998) (see Figure 1.2). It is one of the largest and least stable coagulation factors and has a half life of about 12 hours in adults (shorter in children) (Bolton-Maggs and Pasi 2003).

Figure 1.2: The lifespan of FVIII (figure adapted from Lenting, van Mourik et al. 1998)

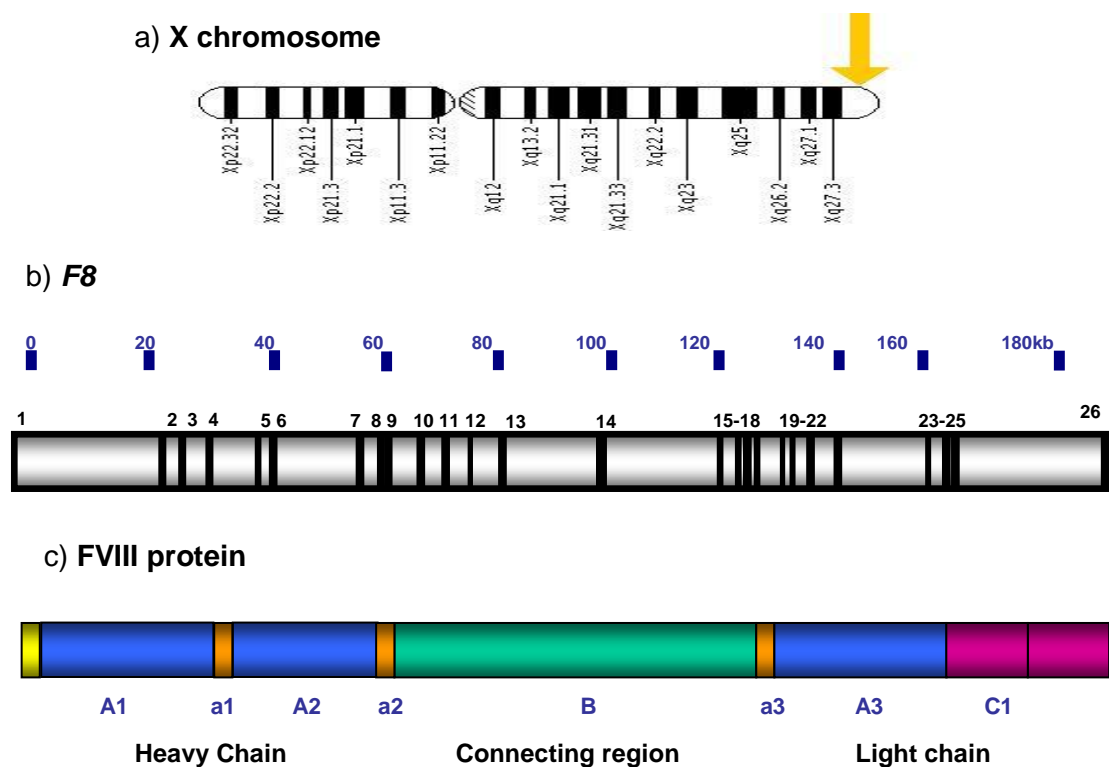


1.2.3. Molecular Biology of FVIII

The molecular characterisation of the human FVIII gene (*F8*) was first carried out in 1984, where it was found to be the largest gene that had been characterised by that time, spanning nearly 0.1% (186 kb) of the X-

chromosome. It has been mapped to the long arm of X-chromosome at Xq28, and consists of 26 exons which range in size from 69 to 3,106 base pairs (bp) (Gitschier, Wood et al. 1984) and encode a polypeptide chain of 2351 amino acids (aa) (see Figure 1.3a&b). This includes a signal peptide of 19 and a mature protein of 2332 aa. The primary structure of FVIII has been shown to consist of a number of domains (see Figure 1.3c). The A domains show approximately 30% homology to one another and display a comparable level of homology to the copper-binding protein ceruloplasmin and to FV, the cofactor in the prothrombinase complex. These domains are surrounded by short spacers (a1, a2 and a3) that contain clusters of aspartic acid and glutamic acid residues, known as acidic regions. The C domains are structurally associated with the C domains of FV. The large B domain does not exhibit any major homology with any other known protein, its function is unknown and is not required for coagulant activity (Lenting, van Mourik et al. 1998).

Figure 1.3: Schematic representation of the chromosomal localisation (a), exonic structure of *F8* (b) and protein domains (c).



1.3 VON WILLEBRAND FACTOR

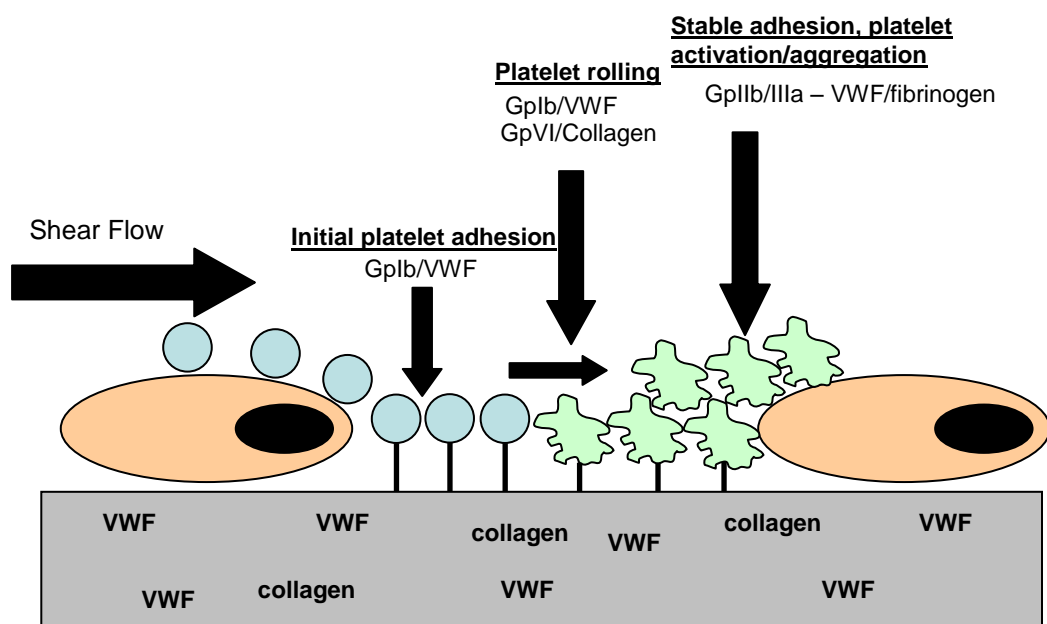
VWF is a multimeric adhesive glycoprotein that is found in plasma, platelet α -granules and subendothelial connective tissue. Plasma VWF plays two essential roles in primary haemostasis as a carrier protein for FVIII, and a promoter of platelet adhesion to the damaged vessel wall (Hoffbrand 2005).

1.3.1 VWF function in haemostasis and coagulation

The main purpose of VWF is to promote thrombus formation by aiding adhesion of platelets to each other and the damaged vessel wall. The circulating pool of VWF is essential for initiating platelet adhesion as it is not found in many vessels where platelet function is required for haemostasis (Bahnak, Wu et al. 1989). VWF is especially important in small blood vessels where there is high shear stress due to rapid blood flow and platelet adhesion and normal primary haemostasis rely on its adhesive properties (Savage, Saldivar et al. 1996).

The initial step in platelet interaction to exposed subendothelium is the binding of the VWF A1 domain to glycoprotein Ib (GpIb) which tethers the platelets to the vessel wall whilst they roll on the surface. The platelets then become activated and VWF binds to the glycoprotein IIb/IIIa (GpIIb/IIIa) complex in an irreversible manner to consolidate the platelet plug. Fibrinogen and thrombin formation are required but not crucial to the process (Ruggeri 1997). The stable adhesion is reinforced by GpVI which is exposed on the platelet surface and binds to collagen which is exposed by the subendothelium at the site of endothelial cell injury.

Figure 1.4: Mechanism of Platelet Adhesion to VWF (figure adapted from Robertson, Lillicrap et al. 2008) In high shear stress conditions, platelets roll along the endothelial cell surface and are tethered to the site of injury through the binding of subendothelial VWF to the GpIb protein of the Ib/IX receptor. Subsequently platelets are activated and the GpIIb/IIIa and GpVI are exposed on the platelet surface. GpIIb/IIIa binds to VWF to permanently tether the platelet to the vessel wall and GpVI binds to collagen as an additional anchor. Interaction of fibrinogen and VWF with GpIIb/IIIa then consolidates the platelet adhesive events and initiates platelet aggregation.



VWF also plays a key role in the stabilisation of FVIII in the blood as it forms a non-covalently bound complex with FVIII (VWF-FVIII) (see Figure 1.4). This complex protects FVIII from activated protein C degradation and carries it to sites of platelet plug and subsequent clot formation (Koppelman, van Hoesj et al. 1996). As a result reduced levels of VWF are associated with reduced levels of circulating FVIII, which are often seen in patients with VWD.

1.3.2 Synthesis of VWF

VWF is synthesised within EC's as a polypeptide of 2813 aa (Jaffe, Hoyer et al. 1973) which undergoes initial dimerization and then multimerization up to a

multimer with a molecular weight (MW) of 20 million Daltons. The higher molecular weight multimers are functionally more effective in promoting platelet adhesion and aggregation. The endothelial cells either directly release VWF into the plasma through a constitutive secretory pathway or store it in endothelial Weibel-Palade bodies where it also drives their formation (Michaux, Abbitt et al. 2006). The VWF stored within these Weibel-Palade bodies is thought to be made up of the largest multimeric species which is not usually seen in the blood of normal individuals. When released at sites of injury they present the most thrombogenic forms of VWF exactly where needed. Circulating high molecular weight (HMW) multimers are cleaved by the metalloproteinase, known as ADAMTS-13 (**a** disintegrin **and** metalloproteinase with **thrombos**pondin type 1 motif), which regulates VWF platelet dependent activity by reducing the size of VWF multimers.

VWF is also synthesised in the megakaryocytes (Nachman, Levine et al. 1977) and stored in the platelet α -granules (which may contain as much as 20 % of the total VWF present in blood) and on activation, secreted by the platelet release reaction. The VWF stored within the platelet α -granules consists of the ultra-large VWF multimers and can be released upon stimulation by a number of agonists including ADP, collagen and thrombin. This allows accumulation of VWF at the site of vascular injury where it can promote further platelet adhesion and thus haemostasis. The VWF that is released from storage sites does not have FVIII bound to it, unlike circulating VWF/FVIII complex which enables FVIII to be concentrated at the site of vascular damage prior to subsequent detachment from VWF to serve as a clotting factor (Nesheim, Pittman et al. 1991).

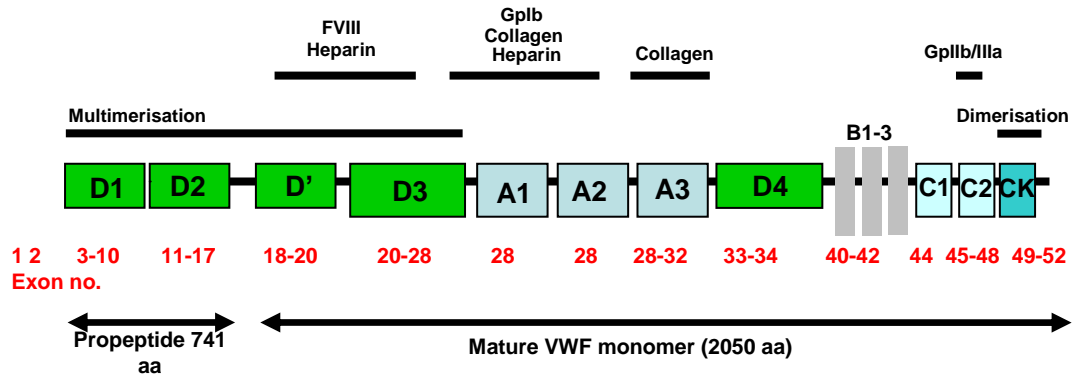
1.3.3. Molecular Biology of VWF

The VWF gene (*VWF*) was first cloned in 1985 and is located on the short arm of chromosome 12 at the locus 12p13.2. It is a large gene spanning 178 kb and is composed of 52 exons (Ginsburg, Handin et al. 1985). *VWF* transcribes a messenger ribonucleic acid (mRNA) product of 8.7 kb. This is translated to produce a 2813 aa product which consists of a signal peptide of 22 aa, a large propeptide of 741 aa and a mature VWF molecule containing 2050 aa

(Bonthron, Orr et al. 1986; Verweij, Diergaarde et al. 1986). A non-processed VWF partial pseudogene (*VWFP*) is located on chromosome 22 which duplicates the region spanning exons 23 through 34. The pseudogene has 97% homology with the genuine *VWF* which may complicate *VWF* analyses upon examination of genomic deoxyribonucleic acid (DNA) (Mancuso, Tuley et al. 1991). The partial pseudogene does not produce a functional transcript due to the presence of splice site and nonsense mutations.

The mature VWF protein of 2050 aa is composed of a sequence of four classes of homologous domains (D, A, B and C) which have specific functions (see Figure 1.5). The domains D1, D2, D' and D3 are involved in multimerisation of the VWF dimers, which altogether span exons 3-28 of *VWF* (Verweij, Hart et al. 1988; Wise, Pittman et al. 1988; Voorberg, Fontijn et al. 1990; Journet, Saffaripour et al. 1993). These domains are rich in cysteine residues which are vital for the multimerisation process. The D' domain displays a binding site for FVIII and heparin as well as possible binding sites for P-selectin which is capable of anchoring ultra-large VWF to the surface of EC's, thus introducing VWF cleavage sites to ADAMTS-13 (Padilla, Moake et al. 2004). The A1 domain exhibits binding sites for the platelet receptor GPIIb α , and contains supplementary sites for heparin, sulphated glycolipids, the snake venom botrocetin and collagen type VI (Ruggeri 2003). Mutations present in this region will cause the molecule to be either more reactive (gain of function mutation, type 2B) or less reactive (type 2A or 2M). The A3 domain is the binding region for collagens type I and III. The C1 domain with its Arginine-Glycine-Aspartate (RGD) sequence is the binding site for integrin α IIb β 3 (GPIIb/IIIa). The A2 domain contains the cleavage site for ADAMTS-13. The cysteine knot (CK) domain spanning exons 49 to 52, is rich in cysteine residues which are essential for the dimerisation of VWF monomers (Voorberg, Fontijn et al. 1991). The domains D4, B1, B2 and B3 which cover exons 33-44 do not yet have a known role, however they may play an important part in the structure of the VWF protein.

Figure 1.5: Schematic representation of the functional domains and binding sites of the VWF Protein.



1.4 HAEMOPHILIA A

Haemophilia A is the most common X-linked inherited bleeding disorder affecting about 1 in 5000 males worldwide. The disorder is caused by low concentrations of FVIII and can be diagnosed by measuring plasma levels of FVIII (FVIII:C). The clinical severity is closely related to the level of activity of FVIII in the plasma and can be classified as mild, moderate or severe, depending on the level. Severely affected haemophiliacs have plasma levels below 1 IU/dL (<1% of normal) and most frequently suffer from serious muscle and joint bleeds. Individuals who are moderately affected have plasma levels of between 1 and 5 IU/dL (1-5% of normal) and those with mild disease have levels of between 5 and 40 IU/dL (>5 to <40% of normal) (White, Rosendaal et al. 2001).

Most cases are referred for investigation due to family history of the disease, however approximately one third of cases occur sporadically. Severe haemophiliacs usually present at about 1 year of age when they become active with unexplained bruising and bleeding. Earlier presentation may be seen in those who have received an intramuscular injection or experienced surgery or trauma. Spontaneous bleeding is also common, occurring once or twice weekly. A diagnosis is usually made following measurement of their FVIII:C. In

males where there is a family history of haemophilia, the cord blood can be tested at birth for FVIII:C. Those with mild or moderate disease may not present until they are adults when they are likely to be exposed to greater trauma (Keeney, Mitchell et al. 2005).

The most common sites for bleeds are the main load-bearing joints i.e. knees, elbows, ankles, hips and shoulders. The bleeding from the synovium into the joint cavity causes severe swelling, stiffness, pain and inflammation. As blood is highly irritating to synovium, this initiates a strong inflammatory reaction which leads to synovial overgrowth and friable vascular tissue, which predisposes it to further haemorrhage. If left untreated, the joint will be considerably weakened by the repeated bleeds and secondary osteoarthritis can occur, leading to physical disability. The large weight bearing muscles such as the iliopsoas, calf, gluteal and forearm muscles can also be affected by bleeding leading to impairment of the blood supply, tissue ischemia, necrosis and subsequent fibrosis and muscle shortening (Hoffbrand 2005).

1.4.1 Laboratory diagnosis of Haemophilia A

In the phenotypic diagnosis of haemophilia A, an initial screening test may show a prolonged activated partial thromboplastin time (APTT), normal prothrombin time (PT), thrombin clotting time (TCT) and bleeding time (BT) with a normal platelet count. Specific assays using factor deficient plasma can be used to detect and quantitate a deficiency. Once it has been established that FVIII is involved the FVIII:C assay is performed. This consists of measuring the intrinsic and common pathways of the coagulation cascade, with FVIII deficient plasma and varying dilutions of the patient plasma. Correction of the clotting time of the deficient plasma by the patient plasma indicates there is no deficiency, however if the times are not corrected then the patient has a FVIII deficiency. The results are compared to reference plasma containing known factor levels and the deficient FVIII level in the patient plasma can be quantitated (showing FVIII clotting activity below 50 IU/dL). Assays for VWF antigen (VWF:Ag) and ristocetin cofactor activity (VWF:RCo) display normal results and it is important that these are carried out in order to distinguish FVIII deficiency due to haemophilia A from VWD (see section 1.13).

Prior to the advent of routine genetic analysis, attempts were made to diagnose the carrier status of haemophilia A in females by phenotypic tests. The routine clotting assay FVIII:C can be used to measure FVIII activity together with FVIII antigen (FVIII:Ag) as the defective gene is the one responsible for FVIII production. VWF antigen levels can also be estimated by immunoassay although these should be normal as VWF production is autosomally controlled and is unaffected in haemophilia. In the normal population the range of values for these factors is between 50 and 200 IU/dL. The FVIII:C level in carriers with the haemophilia defect will be lower and average levels of about 50 IU/dL are found. The ratio of FVIII:C/VWF:Ag can be a useful indicator, averaging at 1.0 in normal individuals and 0.5 in carriers. However the effect of lyonisation (skewed inactivation of the X-chromosome) can result in carriers with markedly decreased FVIII:C levels and some with normal FVIII:C, making carrier detection in the latter group extremely difficult. There are approximately 10% of carriers who will appear phenotypically normal (Lillicrap, White et al. 1987). Therefore, a diagnosis should not be based on these results alone as there are a number of factors affecting FVIII:C and VWF:Ag levels which can hinder the ability to distinguish a minority of carriers from normal females which include ABO blood group, age and pregnancy (Graham, Rizza et al. 1986). To improve the accuracy of phenotypic carrier diagnosis, discriminant function analysis can be used which takes these factors into account as well as FVIII:C and VWF:Ag levels. Phenotypic tests for carrier diagnosis have now been effectively superseded by genetic analysis although they may still be used in developing countries.

1.4.2 Treatment of Haemophilia A

As there is no cure for haemophilia, treatment is based upon raising the concentration of FVIII to arrest haemorrhage or to cover surgery. In some mild haemophiliacs, the level of FVIII can be raised transiently by the use of desmopressin (DDAVP) which is an analogue of antidiuretic hormone. In all other cases, a concentrate of FVIII should be administered. Plasma-derived and recombinant FVIII concentrates are available. Recombinant products are the preferred treatment if available and affordable as they remove the risk of transmission of human and animal infectious agents (Bolton-Maggs and Pasi

2003). Prophylactic treatment with FVIII concentrates can be given regularly to prevent spontaneous joint bleeds from occurring. Prophylaxis is the recommended strategy for children with severe haemophilia and should begin by the age of 4 years, as by this time most affected children will have experienced their first joint bleed. The initial cost of prophylaxis is expensive due to the products, however over the lifetime of an adult with normal joints and good quality of life it proves to be cost effective (Bolton-Maggs and Pasi 2003).

1.4.3 Differential Diagnosis of Haemophilia A

1.4.3.1 Type 2N VWD

Specific variants of coagulation proteins that directly interact with FVIII such as VWF, FII, FIX and FX may affect the amount and function of FVIII. In some cases mutations in *VWF* associated with the VWD variant type 2N, have an impact on the level of FVIII in the circulation and are responsible for the bleeding phenotype seen in these patients, indistinguishable from the one dependent upon mutations in *F8* itself. Both disorders display reduced levels of FVIII:C (~5-40 IU/dL) and normal-to-borderline-low levels of VWF. The symptoms of type 2N VWD are similar to those of mild haemophilia A and include excessive bleeding during surgery or other invasive procedures. The best way to distinguish between the two disorders is to use the VWF:FVIII binding test (VWF:FVIII B), which determines the ability of VWF to bind FVIII (Casonato, Pontara et al. 2007). Type 2N VWD should be suspected in patients diagnosed with mild haemophilia A who either do not respond well to FVIII infusions or belong to families in which males and females are equally affected. Missense mutations in *VWF* reduce the ability of VWF to bind and protect FVIII, therefore VWF and FVIII levels can look identical to those in males with mild haemophilia A or in symptomatic haemophilia A carrier females.

1.4.3.2 Combined Deficiency OF FV and FVIII (F5F8D)

The combined deficiency of coagulation FV and FVIII (F5F8D) is an autosomal recessive bleeding disorder. F5F8D is distinguished by coexisting low levels (usually between 5 and 30 IU/dL) of both FV and FVIII and exhibits a mild to moderate bleeding tendency (Zhang and Ginsburg 2004). However F5F8D is

completely independent of FV deficiency and FVIII deficiency and is unusual because it arises because of mutations in two genes (*LMAN-1* and *MCFD2*) which encode proteins involved in the FV and FVIII intracellular transport rather than DNA defects in the genes which encode the corresponding coagulation factors. Null mutations in the *LMAN-1* gene (Lectin Mannose Binding Protein) have been found in approximately 70% of affected patients, while null and missense mutations in *MCFD2* (Multiple Coagulation Factor Deficiency 2) account for the remaining 30% (Zhang, McGee et al. 2006). Bleeding episodes are treated by replacing FV through fresh frozen plasma and FVIII by DDAVP or specific FVIII concentrates. While F5F8D is extremely rare (1:1,000,000 in the general population), it has an increased frequency in populations where there is a high incidence of consanguineous marriages and therefore should be considered in some cases of mild haemophilia A where no mutation is detectable.

1.5 MOLECULAR PATHOLOGY OF HAEMOPHILIA A

The molecular basis of haemophilia A is extremely diverse and there are approximately 2100 individual mutations which have been found in *F8* which have been attributed to the clinical phenotype. The most common gene defects seen in approximately 90% of patients are point mutations (single nucleotide substitutions), followed by deletions which are found in approximately 5-10% of patients. Insertions and rearrangements/inversions are rarely seen in haemophilia patients, with the exception of two common inversions involving intron 1 and 22 the intron found in patients with severe haemophilia A (Bowen 2002). The Haemophilia A Mutation Database and FVIII Resource Site (HAMSTeRS) is a locus specific database which lists mutations that have been characterised in haemophilia A and can be readily accessed (www.hadb.org.uk).

This molecular characterisation has also led to developments in the production of recombinant therapeutic clotting-factor concentrates, the generation of animal disease models and the production of wild-type and mutant proteins for structure-function analysis (Bolton-Maggs and Pasi 2003).

1.5.1 Gross DNA Rearrangements

1.5.1.1 Common Partial Inversions of FVIII

The intron 22 inversion mutation is responsible for disease in 20% of all patients and always results in severe disease. It is the causative mutation in nearly 45% of severe haemophilia A cases (Lakich, Kazazian et al. 1993; Naylor, Brinke et al. 1993). The intron 22 inversion occurs because of homologous recombination between copies of a repeated DNA sequence, the intron 22 homologous region (int22h), one copy located in intron 22 of FVIII, the other two copies distal and telomeric to FVIII. It is primarily an error of DNA replication during spermatogenesis and completely disrupts *F8* leading to severe disease. Until relatively recently the intron 22 inversion was detected by Southern blot analysis, however this has been replaced by long polymerase chain reaction (PCR) protocols which are able to detect normal and inverted genes (Liu, Nozari et al. 1998) and more recently still by inverse PCR protocols (Rossetti, Radic et al. 2005; Rossetti, Radic et al. 2008).

The intron 1 inversion is the second most common mutation in severe haemophilia A and has been reported to have a frequency of 1.8% in the UK population (Cumming 2004). The mutation can be detected by a robust dual PCR assay which has been designed to amplify sequence flanking the intron 1 homologous (int1h) regions independently (Bagnall, Waseem et al. 2002).

1.5.1.2 Large Deletions

Deletions affecting *F8* include whole gene deletions, partial gene deletions at the 5' end, 3' end, or within the gene, and microdeletions of one to several base pairs. They do not seem to be confined to a specific region of the gene but appear to be randomly distributed throughout. Deletions have the capability of damaging genetic function, removing amino acids, or introducing a frameshift which is why they are generally associated with severe disease.

Large deletions (more than 100 nucleotides) are found in approximately 5% of patients with haemophilia A (Antonarakis 1998). As of November 2012, the HAMSTeRS database contains 255 large deletion mutations which range from less than 1 kb up to more than 210 kb deleting the entire gene, which is

probably the result of non-homologous recombination (Woods-Samuels, Kazazian et al. 1991). Large deletions in *F8* are normally associated with clinically severe disease with no measurable FVIII activity in plasma samples and undetectable antigen as any truncated proteins produced are either poorly expressed, inactive or cleared from the circulation. There are known reports of a deletion of exon 22 and also of exons 23-24 that are associated with moderate disease. This may result from in-frame splicing of mRNA leading to the production of hypoactive FVIII which lacks short stretches of amino acids (Youssoufian, Antonarakis et al. 1987; Lavergne, Bahnak et al. 1992). Exon-spanning deletions are easily detected in affected males due to lack of PCR amplification of missing exons of *F8*. However, in carrier females, these mutations are masked by amplification of the normal allele. For methods used to identify female carrier status see section 1.9.

1.5.1.3 Insertions of Retrotransposons

A very small number of patients have an insertion of a repetitive element (LINE1 or Alu) which causes severe haemophilia by disrupting the coding sequence of the gene. The de novo insertion of long interspersed nuclear elements (LINEs) in the human genome was first reported in *F8* in two severe haemophilia A patients (Kazazian, Wong et al. 1988). The first case was a 3.8 kb portion of a LINE element inserted in exon 14 of FVIII. The inserted DNA had a poly(A) tail, produced a target site duplication, and was inserted in a relatively adenine-rich sequence of exon 14. The second case involved the insertion of a 2.1 kb portion of a LINE element in a different section of exon 14 and had all the features of a retrotransposition. LINE insertions or other retrotransposons are relatively uncommon as there has only been two examples in >1000 patients studied.

1.5.1.4 Duplications

There have been several cases reported involving duplications in *F8*. The first report described a duplication of 23 kb of intron 22 inserted between exons 23 and 25 (Gitschier 1988). This rearrangement found in two female siblings was apparently unstable and led to deletion of exons 23-25 in the male offspring of one of the females. A second exon duplication of *F8* was identified in 1990 and

was characterised as an in-frame tandem duplication of exon 13 that produces an unstable FVIII protein with a short half-life (Murru, Casula et al. 1990). This has since been shown to be present in up to one third of cases in the northern Italian mild haemophilia population (Acquila, Pasino et al. 2004).

As standard PCR-based techniques do not indicate relative template dosage, further *F8* duplications have not been reported until recently following the introduction of Multiplex Ligation-dependent Probe Amplification (MLPA) (see section 1.9.1). A study carried out by Rost, Loffler et al. (2008) reanalysed 80 mutation-negative haemophilia A patients from a cohort of about 2000 using MLPA and identified exon duplications in nine cases, a frequency of ~0.5% (see section 1.10).

1.5.2 Point Mutations

The point mutations that arise in haemophilia are made up of missense mutations (where a codon is changed so that a different amino acid is encoded), nonsense point mutations (where an amino acid codon is changed into a premature translation stop codon) and mRNA splice site point mutations (which corrupt a true mRNA splice site, or create a novel one). The HAMSTeRS database last updated in November 2012 reported a total of 1349 unique single base (point) mutations, of which 983 were missense, 208 were nonsense and 158 were splice site related.

1.5.2.1 Missense Mutations

Missense point mutations can give rise to varying severity of haemophilia depending on the nature and location of the amino acid substitution. Less severe disease is associated with semi-conservative amino acid substitutions where the new amino acid has very similar properties to the normal one, unless they occur in regions of structural and functional importance. Whereas moderate to severe disease tends to arise because of non-conservative amino acid substitutions, no matter where they occur in the protein. With non-conservative amino acid substitutions there is a higher chance of them having an overall more damaging effect as they have the capacity for structural perturbation/destabilisation (Bowen 2002).

1.5.2.2 Nonsense Mutations

Severe haemophilia is associated with nonsense mutations and this is because the resulting transcripts may undergo nonsense mediated ribonucleic acid (RNA) decay or, if translated, result in a truncated protein molecule which is highly likely to lack function. Another negative effect which can arise from nonsense mutations is exon skipping which can cause either an in frame skip leading to a protein lacking the amino acids encoded by the skipped exon, or an out of frame skip which results in a frame shift (Bowen 2002).

1.5.2.3 mRNA Splice Site Mutations

Mutations that destroy or create mRNA splice sites are associated with variable severity of haemophilia, depending on whether some correct transcripts can be processed (mild to moderate disease) or whether there is a complete loss of correct mRNA processing (severe disease). Splice site mutations can also cause exon skipping with the overall severity depending on whether the skip is in frame or results in a frame shift.

1.5.2.4 Small Deletions/Insertions

Small deletions in the coding region of *F8* are generally associated with severe haemophilia A. They vary in size from 1 bp to 86 bp and most result in frame shifts which consequently eliminate FVIII expression. More than one-third of small deletions are found in the largest exon, exon 14 and the majority occur in DNA regions of short direct repeats (Antonarakis 1998).

Insertions have also been characterised in *F8* and like deletions can be gross or as small as one or several nucleotides. They range in size from 1 bp up to 2.1 kb and 3.8 kb LINE elements. Most insertions are of 1 bp, often an A in a stretch of A residues. They are usually associated with severe disease as gene function or the gene product can be adversely affected. However an exception to this is small deletion/insertion mutations within poly-A runs of *F8* which have been shown to be associated with a less severe haemophilia A phenotype. Small deletions/insertions resulting in uninterrupted stretches of adenines are likely to be sensitive to DNA replication/RNA transcription and translation errors thereby promoting the mechanism of partial correction of the reading frame.

Subsequently, affected individuals have a milder than expected form of haemophilia A (Young, Inaba et al. 1997).

1.6 FACTOR VIII INHIBITORS

Approximately 30-50% of patients with severe haemophilia A develop antibodies (known as inhibitors) to FVIII following treatment with exogenous FVIII which is due to insufficient synthesis of endogenous FVIII protein by the individual to induce immune tolerance. These inhibitors may be either low titre antibodies, managed by increased or continuing treatment with FVIII concentrates, or high titre antibodies which prevent treatment with factor concentrates (Bolton-Maggs and Pasi 2003).

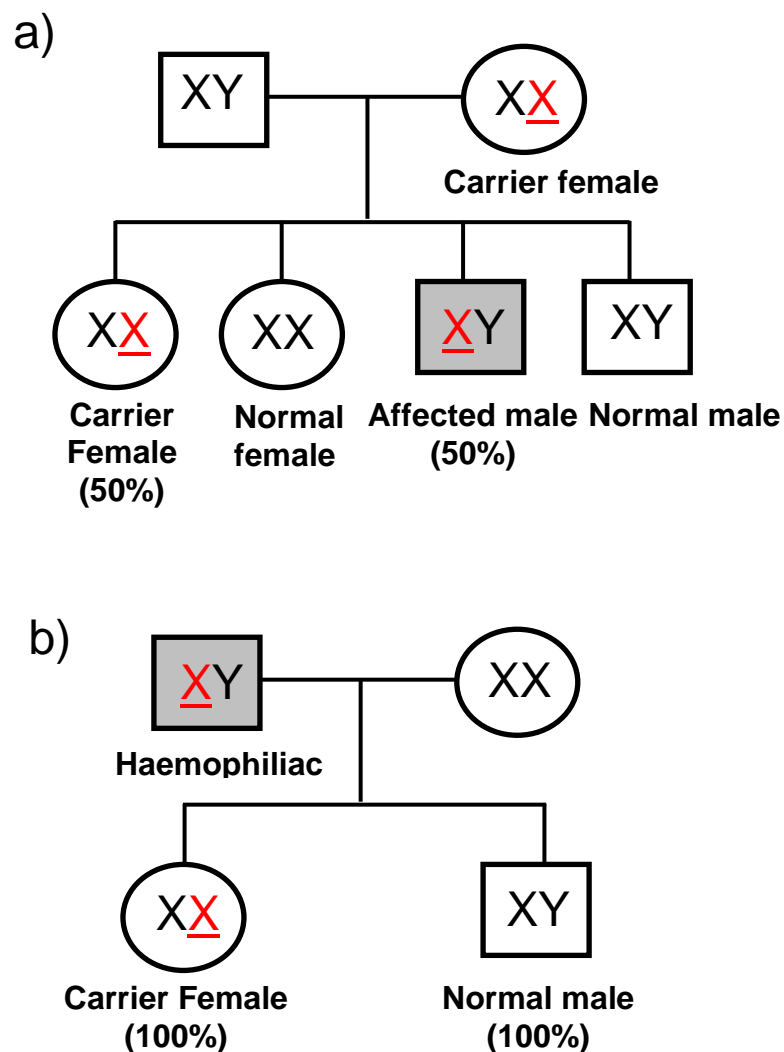
The risk of developing an inhibitor is related to the nature of the molecular defect in *F8*. Nonsense mutations, deletions and inversions in *F8* are most commonly found in the majority of reported inhibitor cases. Patients with these types of mutation have a 35% higher incidence of developing an inhibitor compared to those carrying missense mutations or small deletions because they can either severely truncate or stop production of the protein. Gross deletions of *F8* have also been found to cause a two- to three-fold increased incidence of inhibitors than for patients without detectable deletions (Millar, Steinbrecher et al. 1990).

1.7 INHERITANCE OF HAEMOPHILIA A

A female carrier of the defective X-linked gene has a 50% chance of passing it on to her children, resulting in affected males or carrier females (see Figure 1.6a). The daughters of a haemophiliac father will be obligatory carriers of the defective gene (see Figure 1.6b). Sporadic haemophilia is caused by *de novo* mutations and represents approximately one third of cases. Females at risk of being a carrier can be identified by validating the pattern of inheritance by genetic analysis which is the definitive diagnostic tool for carriership analysis. In families with sporadic cases, the risk of somatic mosaicism must be considered as it causes uncertainty about the recurrence risk in parents who appear to be non-carriers (Peyvandi, Jayandharan et al. 2006). Severe haemophilia in females is rare, however it can arise from genetic mechanisms

such as extreme lyonisation, Turner's syndrome or transmission of a mutation by both parents (father with haemophilia and mother a carrier). Following haemophilia diagnosis, an accurate family history must be taken in order to create a pedigree by which other family members at risk of passing it on to their offspring can be counselled.

Figure 1.6: Inheritance in haemophilia (figure adapted from Peyvandi, Jayandharan et al. 2006). The X indicates the defective gene that is inherited.



1.8 GENETIC DIAGNOSIS OF HAEMOPHILIA A

The practice guidelines for the molecular diagnosis of haemophilia A recommend that all children with haemophilia in the UK are investigated to

establish the causative *F8* mutation (UKHCDO 2010). Genetic analysis is also necessary for the reliable identification of female carriers following phenotypic testing, therefore allowing them to make informed choices should they wish to start a family or in the early stages of pregnancy. Genetic counselling should always be carried out by suitably qualified health professionals who have a detailed knowledge of haemophilia. All testing should be performed in a member laboratory of the UK Haemophilia Centre Doctors' Organisation Haemophilia Genetics Laboratory Network (UKHCDO-GLN) working to agreed peer-reviewed standards of quality (Keeney, Mitchell et al. 2005).

The genetic diagnostic strategy is influenced by the severity of the disorder within a family therefore this should be determined before any screening is undertaken. Severe haemophiliacs should be screened for the *F8* intron 22 inversion mutation followed by the *F8* intron 1 inversion mutation, which will identify the underlying mutation in 45-50% of cases. For the remaining pedigrees, full mutational analysis of *F8* should be performed. Mild/moderate haemophiliacs require full mutation analysis as they are not associated with a common mutational mechanism (UKHCDO 2010).

Historically, the most common method for resolving female carrier status was linkage analysis; however this has now been overtaken by direct mutation detection, including DNA sequencing, which is considered the gold standard for mutation detection. Linkage analysis may also fail in a number of families due to either lack of prior family history, key pedigree members not being available, polymorphisms which are uninformative in key females or non-paternity. In the UK, mutations are usually identified by direct DNA sequencing of *F8* in affected males and then confirmed or excluded in female relatives. Intragenic linked markers may still be useful in circumstances where a family has previously been investigated by linked markers and the mutation has not been identified, or where a mutation has not been verified or found, although where a mutation has not been identified in an index male after full mutation analysis has been performed then linkage of the FVIII deficiency to *F8* cannot be assumed.

1.8 MOSAICISM

Genetic diagnosis in haemophilia A can be complicated by germ line and somatic mosaicism which has been reported in 13% of families with sporadic severe haemophilia associated with point mutations (Leuer, Oldenburg et al. 2001). Therefore the prospect of mosaicism should be considered in sporadic haemophilia where the mother of an affected male does not seem to carry the mutation in her leucocyte DNA, especially when the *de novo* mutation appears to be a point mutation.

1.9 DETECTION OF LARGE SCALE DELETIONS IN HETEROZYGOUS CARRIERS

Large scale mutations are easily detected in affected males as the missing regions of *F8* fail to amplify. However heterozygous carriers with these mutations may not be identified due to amplification of one normal allele. Methods that can be applied to the *F8* locus to establish female carrier status in these families include; linkage analysis, gene dosage analysis, gap-PCR (where deletion boundaries are already known) and other indirect methods such as Multiplex Ligation-dependent Probe Amplification (MLPA) (Keeney, Mitchell et al. 2005).

1.9.1. MULTIPLEX LIGATION-DEPENDENT PROBE AMPLIFICATION (MLPA)

MLPA is a fairly new multiplex PCR method which is used to quantify the relative copy number of up to 50 different genomic DNA or RNA sequences of a gene in one single reaction (Schouten, McElgunn et al. 2002). The MLPA reaction consists of denatured genomic DNA which is hybridised with region or exon-specific probes. Each probe is made up of 2 oligonucleotides, one synthetic and one M13-derived which hybridise to adjacent target sequences and are then ligated. The probe ligation products are subsequently amplified by PCR using only one primer pair. The quantity of amplification product is a measure for the number of total target sequences in the sample. The amplification products are separated by capillary electrophoresis and compared to a control DNA sample, to determine the relative copy number of the target.

The use of MLPA for detecting copy numbers offers many advantages compared to other techniques. Methods such as sequencing fail to detect copy number changes, while Southern Blot analysis which is capable of detecting many aberrations is not ideal as a routine technique due to its laborious, technically demanding nature which has been largely superseded by PCR based methods where possible. When compared to florescent in situ hybridisation (FISH), MLPA is more sensitive at detecting some single gene aberrations. It can be used on purified DNA and is a relatively simple and fairly robust method. Probe sets are now commercially available from MRC Holland (P178-B1 *F8*) for the *F8* locus capable of detecting deletions/duplications of one or more exons. In the case of *F8*, the deletion of probe recognition sequences in affected males will be obvious by the absence of the probe amplification product. In female carriers, you would expect to see a 35-50% reduced relative peak area of the amplification product of a probe. For full explanation of the MLPA technique see Materials and methods section 2.4.

1.10 MOLECULAR STUDIES OF LARGE SCALE DELETIONS / DUPLICATIONS IN HAEMOPHILIA

Various studies have been recently carried out to determine the underlying genetic defect in patients where it was not clear. It is still true that the intron 22 and intron 1 inversion mutations are the most common causative mutation in severe haemophilia and in the remainder of cases, the molecular defects are predominantly single base-pair substitutions, but large deletions or insertions are found in approximately 5% of patients (Becker, Schwaab et al. 1996).

Pavlova, Forster et al. (2008) reported an assay based on multiplex amplification of a number of exons of *F8* of differing lengths which they consequently quantitated by liquid chromatography (LC). Using this method they investigated 19 families with severe haemophilia A where an index patient exhibited a large deletion of one or more exons. Of 24 possible carriers they were able to accurately classify 16 female carriers and eight non-carriers for the presence of a large deletion. One mother and one grandmother were determined to be non-carriers. This study high-lighted the high *de novo* mutation rate of large deletions in female germ cells. They concluded that the

multiplex PCR-LC technique has been shown to be a rapid, simple and reliable method for detection of heterozygous large deletions in carrier females.

A study by Rost, Loffler et al. (2008) investigated 80 mutation negative cases comprising severe, moderate and mild haemophiliacs from a cohort of 2000 haemophilia A patients. The 80 patients were screened using MLPA analysis and among these patients they detected nine large duplications. Two cases showed duplications of single exons 13 and 14, while the other duplications affected more than one exon. In the patients studied they discovered seven novel large deletions and one recently described single-exon duplication within *F8*. Only one patient with mild haemophilia A was found to have an exon 13 duplication which leads to an internal in frame duplication within the FVIII protein. This mutation has been shown to be present in 32% of northern Italian patients with mild haemophilia A, likely due to a founder effect. The mother of this single patient was also found to have an exon 13 peak with a ratio of 1.5 by MLPA which identified her as carrier for the duplication and hence haemophilia A. The remaining eight patients all had severe haemophilia A with less than 1% FVIII activity. All the novel duplications found were expected to produce frame shifts of the *F8* mRNA and protein sequence if they were re-arranged in a head-to-tail mode (tandem duplication) which is in concurrence with the severe phenotype of these patients. The development of inhibitors did not occur in seven of the severe haemophiliacs.

Not all large duplications occur in tandem and may result in transposition to a genomic position outside of the affected gene locus. However the haemophilia phenotype is unlikely to be caused by this type of mechanism as it should not influence the transcription and translation of the intact *F8*. In conclusion from the cohort studied, approximately 0.5% of 2000 haemophilia A patients showed large duplications of one or more exons which add to the recurrent low-frequency mutation types like the inversion of intron 1.

Another study by Zimmermann, Oldenburg et al. (2010) was the first to carry out a broad analysis of sequences that produced exon-spanning duplications affecting *F8* using MLPA. They identified breakpoints in 10 out of 16 unrelated

cases with large duplications, including those identified by Rost, Loffler et al. (2008). These involved nine tandem duplications plus one mixed duplication (exons 1-22) and triplication (exons 23-25), affecting regions throughout *F8* and ranging from the duplication of exon 6 alone to duplication of exons 2-25. Characterisation of these duplications showed that LINEs and SINEs are often found in proximity to *F8* duplication breakpoints. Three mutations with Alu SINEs at each breakpoint shared significant sequence homology (>78%) while two mutations with LINEs at each breakpoint illustrated only slight flanking-sequence homology. These events could originate from mechanisms such as homologous and illegitimate recombination. Other more complex rearrangements could also be responsible as six mutations were unable to be characterised due to unsuccessful amplification. Most introns involved in breakpoints are the same as those already associated with large *F8* deletions, for example introns 13, 22 and 25. This is due to the fact that some of these regions have an abundance of repetitive sequence elements which present sequences that are prone to recombination.

Theophilus, Baugh et al. (2011) investigated 11 families among a cohort of about 400 in whom a mutation could not be identified by traditional screening methods, using MLPA analysis. They identified whole exon deletions or duplications in seven families affected with severe haemophilia, but not in one family with moderate or the three families with mild haemophilia A. In four of the severe patients, they suspected that deletions of 1 to 6 exons of *F8* were responsible due to repeated failure to amplify the corresponding exons by PCR. Using MLPA they were able to confirm deletions of exons 1, 1-6, 6 and 25 in these severe families which have all been previously associated with a severe phenotype. In two cases, they were also able to confirm the gene abnormality in the affected male's carrier mother. In the remaining three families with severe haemophilia A, they identified novel duplications of exons 2-12 and exon 15 and a previously reported duplication of exons 14-22. These duplications would result in a frame shift if located in tandem orientation. Rost, Loffler et al. (2008) have previously reported duplications of exons 14, 1-5, 5-25, 23-25, 2-25, 14-21 and 7-11 and one recently described duplication of exon 13 within *F8*.

In the remaining families with mild and moderate haemophilia, MLPA analysis demonstrated normal gene dosage results for all 26 *F8* exons. They had been fully analysed by DNA sequencing and combined F5F8D and type 2N VWD had been ruled out. Therefore they speculated that the underlying mutations responsible for their disease may be located deep within introns or comprise undiscovered mutation types in *F8* that are undetectable by current methods. On the other hand the reduced FVIII:C levels may be due to mutations in other genes which affect *F8* expression.

Theophilus, Baugh et al. (2011) also looked into the inhibitor status for each index case and they found that only one index case with a deletion encompassing the promoter region to exon 6) had high titre non-transient inhibitors. Interestingly, in comparison, a study carried out in 100 Korean patients with severe haemophilia A found that when inhibitor status was correlated with the genotype, the highest risk (100%) was observed in large deletion mutations (4/4). This is consistent with the opinion that large deletions are the single most significant inhibitor risk factor in severe haemophilia A (Kim, Chung et al. 2012).

Another recent study by Rafati, Ravanbod et al. (2011) found ten large deletions and one duplication in *F8* of eleven unrelated Iranian severe haemophilia A families using MLPA. Their results illustrated that 3.2% (11/347) of their severe haemophilia A patients have large deletions or duplications in their *F8* among whom 8 patients developed high-titre antibodies to FVIII.

A report by Lannoy, Abinet et al. (2009) illustrated the importance of genetic analysis of *F8* in female patients presenting with FVIII deficiency in the absence of a family history of haemophilia or VWD. They identified two *F8 de novo* deletions which had not been previously reported, using MLPA, in two girls with isolated factor VIII deficiency. Both deletions were predicted to result in a severe form of haemophilia. The slight reduction in FVIII levels seen in the two girls is attributed to the preferential inactivation of the X chromosome, which carries no deletion. As the patients had no family history of haemophilia and no mutations were identified in both mothers, they concluded that the deletions

were caused by *de novo* mutations occurring in the male germ cells. Therefore both patients were deemed to be carriers of a severe form of haemophilia A and the risk of having an affected son 1:2. The risk of an affected male offspring developing inhibitors to FVIII was estimated at 40%. They also concluded that in cases such as these, MLPA should be employed to exclude *F8* deletions in the absence of the common inversions involving intron 22 or 1 or point mutations.

These studies have shown that the introduction of copy number variation (CNV) analysis, in particular MLPA, are essential for determining haemophilia A carrier-status in female relatives of males with large *F8* deletions. It can also be used to confirm previously suspected exon deletions in males with haemophilia A. MLPA can also be used to successfully detect *F8* duplications following DNA sequence analysis due to the fairly low frequency of duplication mutations. MLPA is an important addition to the repertoire of tests used to detect *F8* mutations and can be implemented after traditional methods or where the latter has failed to identify a causative mutation.

1.11 VON WILLEBRAND DISEASE (VWD)

VWD is the most commonly inherited bleeding disorder in humans with a prevalence of 1:100-1:1000 (Rodeghiero, Castaman et al. 1987; Holmberg and Nilsson 1992). It was first described in 1926 by Erik von Willebrand who discovered a novel bleeding disorder in a large family from Foglo, an island in the Åland archipelago of Finland. He reported a serious haemorrhagic condition which he called 'pseudohaemophilia' because of the autosomal dominant inheritance and attributed to a platelet or vessel wall defect (von Willebrand 1926; von Willebrand 1931).

VWD is now defined as a deficiency of VWF function leading to impaired haemostasis, caused by either a decrease in VWF levels (quantitative deficiency) or a decrease in VWF function (qualitative deficiency). The pattern of inheritance is autosomal dominant or recessive, however women with milder VWD forms appear more symptomatic than men.

1.11.1 Clinical Manifestations

As described in section 1.3, VWF is essential for primary haemostasis, therefore individuals with VWD present with bleeding symptoms such as bruising, epistaxis, oral cavity bleeding, menorrhagia, gastrointestinal bleeding and bleeding after trauma or surgery (Federici 2004). The severity of clinical symptoms varies in individuals as they are dependent upon a number of molecular and genetic factors. Most patients with severe forms of VWD present in childhood whilst milder forms may not present until adulthood when an individual may experience a significant haemostatic challenge. As VWD is such a heterogeneous disorder some patients with milder forms of the disease remain asymptomatic and therefore remain under- or mis-diagnosed (Laffan, Brown et al. 2004). This has led to the classification of VWD based on distinct types and treatment decisions are made according to this. Multiple tests are used to diagnose and classify VWD accurately as no single clinical or laboratory assessment can give a definitive diagnosis.

1.11.2 Bleeding history

The three key components needed to make a diagnosis of VWD are: a positive bleeding history since childhood, a family history of bleeding with an autosomal dominant or recessive pattern of inheritance, and reduced VWF activity in the plasma (Federici 2009). The bleeding severity score (BSS) is a summary score which can be used to quantify the bleeding symptoms in patients suspected of having VWD and is essential in the diagnosis of VWD.

1.12 CLASSIFICATION OF VWD

In 1994 the Subcommittee on VWF published recommendations for the classification of VWD based upon available laboratory tests and clinical characteristics (Sadler 1994). Further research into understanding the pathophysiology of VWD has led to re-evaluation of this classification. The classification now identifies three main categories (see Table 1.1). Type 1 refers to a partial quantitative deficiency of VWF, type 2 represents qualitative abnormalities and type 3 includes a complete deficiency of VWF. Type 2 is further subdivided into four categories based on their phenotype (2A, 2B, 2M, 2N) (Sadler, Budde et al. 2006). These six variants manifest due to different

pathophysiologic mechanisms and are associated with individual clinical features and therapeutic requirements. Although mixed VWF phenotypes can arise due to complex VWF mutations which work alone or in combination.

1.12.1 Type 1 VWD

The most common form of VWD is type 1 which results from a partial, quantitative deficiency of VWF and accounts for 50-75% of all VWD cases seen. A number of factors including reduced secretion, accelerated clearance or enhanced sensitivity to proteolysis can lead to reduced VWF levels. Type 1 VWD diagnosis can often be difficult and the main criterion used is low VWF levels, however this can be influenced by age, sex, exercise and stress and blood group. The significant laboratory findings for type 1 disease include a normal ratio of VWF activity to VWF level in plasma (VWF:RCo/VWF:Ag >0.6), normal multimer size distribution but with a reduced concentration (Ruggeri, Mannucci et al. 1982).

The bleeding diathesis seen in type 1 VWD patients is mild to moderate. Mild type 1 VWD can be harder to diagnose as it is difficult to identify them from individuals with low normal VWF levels due to the broad normal range of plasma VWF (Laffan, Brown et al. 2004).

A definitive diagnosis can be made by autosomal dominant inheritance, abnormal bleeding symptoms and decreased levels of qualitatively normal VWF (Federici 1998; Laffan, Brown et al. 2004). A diagnosis of type 1 VWD can be made difficult by incomplete penetrance and variable expression of the disease (Sadler and Rodeghiero 2005). Several subgroups of VWD type 1 have been identified based on the relative levels of plasma and platelet VWF which indicates that different molecular mechanisms may be responsible. In many affected families it has been seen that a spectrum of mutations rather than one single mutant allele is responsible for disease symptoms. These consist of mainly deletions, frameshifts and nonsense mutations and are similar to that in type 3 VWD (Sadler, Mannucci et al. 2000). Type 1 VWD can also be inherited infrequently as a dominant trait with extremely high penetrance and very low VWF levels. This has been attributed to dominant negative missense

mutations that impair the intracellular transport and secretion of normal VWF subunits (Eikenboom, Matsushita et al. 1996).

1.12.2 Type 2 VWD

Type 2 VWD is due to a qualitative defect of VWF function and is subdivided into four variants (2A, 2B, 2M and 2N) which contribute to 25% of all VWD cases diagnosed (Holmberg and Nilsson 1985).

1.12.2.1 Type 2A VWD

Type 2A VWD is associated with reduced VWF-dependant platelet adhesion and a deficiency of HMW VWF multimers resulting in disproportionate reductions in platelet interaction and collagen binding relative to VWF levels (VWF:RCo/VWF:Ag and VWF:CB/VWF:Ag, respectively). The deficiency of HMW multimers may result from impaired multimer assembly or greater sensitivity to cleavage by ADAMTS-13. The loss of large multimers is associated with unequal reductions in platelet interaction and collagen binding in relation to VWF levels (Favaloro, Lillicrap et al. 2004). The minimum criteria for diagnosis include a personal or family history of bleeding, VWF:RCo below 50 IU/dL and a VWF:Ag level of <0.7 in the absence of HMW multimers. Patients present with a history of mucosal or trauma-related bleeding and inheritance is generally autosomal dominant although some cases may be autosomal recessive (Holmberg and Nilsson 1992).

1.12.2.2 Type 2B VWD

Type 2B VWD is defined as a qualitative gain of function variant of VWF with increased affinity for platelet GpIb. This VWD variant is quite uncommon and accounts for less than 5% of all cases. Patients are diagnosed by a personal or family history of mucosal bleeding, increased sensitivity to RIPA at low-dose ristocetin (≤ 0.75 mg/mL) and exclusion of pseudo (platelet type) VWD. Patients present with a history of mucosal or trauma-related bleeding and inheritance is normally autosomal dominant although recessive cases have been described (Laffan, Brown et al. 2004).

1.12.2.3 Type 2M VWD

Type 2M VWD is a qualitative variant characterised by decreased platelet-dependent function that is not caused by the absence of HMW multimers. Patients present with a history of mucosal or trauma-related bleeding and inheritance is autosomal dominant. A diagnosis can be made by a VWF:RCO below 50 IU/dL and <0.7 of the VWF:Ag level and the presence of HMW multimers (Laffan, Brown et al. 2004). Type 2M is caused by mutations that inactivate specific binding sites for ligands in connective tissue or on the platelet surface without affecting multimer assembly or stability.

1.12.2.4 Type 2N VWD

Type 2N VWD is a recessively inherited disorder with decreased affinity for FVIII. FVIII/VWF binding is essential for normal FVIII survival in the circulation therefore the reduced binding results in accelerated clearance of FVIII. Therefore a diagnosis can be made by decreased FVIII:C level and decreased binding of control FVIII to patient VWF. However this can also cause confusion between a diagnosis of haemophilia A and type 2N VWD and it has been reported that a small number of patients with apparent haemophilia A actually have type 2N VWD (Schneppenheim, Budde et al. 1996). It is important to distinguish VWD type 2N from haemophilia A so that bleeding episodes can be treated appropriately, valid genetic counselling can be given and accurate carrier or pre-natal diagnosis provided (Laffan, Brown et al. 2004).

1.12.3 Type 3 VWD

Type 3 VWD is a severe autosomal recessive disorder which arises because of a quantitative deficiency of the VWF protein. It is characterised by an absence or virtual absence of VWF (<5% of normal) and FVIII activity is notably reduced due to the requirement for VWF to stabilise FVIII in the plasma (Budde and Schneppenheim 2001; Federici 2003). In the general population the incidence of type 3 VWD is reported to be 0.55 to 3.2 per million (Mannucci, Bloom et al. 1984) in Western countries, however a higher prevalence may be seen in populations where consanguinity is widespread (Berliner, Seligsohn et al. 1986; Sadler, Mannucci et al. 2000). Type 3 VWD patients usually present in childhood with symptoms such as severe mucocutaneous haemorrhage and

bleeding following minor trauma (Laffan, Brown et al. 2004). Heterozygous relatives can be seen to have normal or mildly reduced levels of VWF, and are generally asymptomatic although a minority may suffer from mild bleeding symptoms (Schneppenheim, Krey et al. 1994). Both type 3 and type 1 VWD arise from quantitative defects and have a similar pathophysiology, however they have been classified separately as type 3 symptoms are more severe and require a different treatment regime.

Table 1.1: Classification of VWD (adapted from Sadler, Budde et al. 2006).

Type	Description
1	Partial quantitative deficiency of VWF.
2	Qualitative VWF defects.
2A	Decreased VWF-dependant platelet adhesion and a selective deficiency of HMW VWF multimers.
2B	Increased affinity for platelet glycoprotein Ib.
2M	Decreased VWF-dependant platelet adhesion without a selective deficiency of HMW VWF multimers.
2N	Markedly decreased binding affinity for FVIII.
3	Virtually complete deficiency of VWF.

1.13 LABORATORY DIAGNOSIS OF VWD

Laboratory investigations for VWD are only carried out if a diagnosis will be of benefit to advise or treat a patient with bleeding symptoms. The VWF protein can be characterised by a number of different techniques and relies upon these to identify the correct classification of VWD. Primary screening tests are used to initially evaluate a bleeding patient, specific tests are required to confirm a diagnosis of VWD and discriminating tests are essential in distinguishing the subtypes from one another (Sadler, Mannucci et al. 2000).

1.13.1 SCREENING TESTS

1.13.1.1 Bleeding Time (BT)

The BT is the length of time that it takes for a standardised skin wound to stop bleeding. The BT is always prolonged in severe VWD (types 3 and some forms

of type 2). In milder forms of VWD the BT is usually found to be normal or minimally prolonged but can vary in a particular patient. Limitations of this test include insensitivity, lack of specificity and poor reproducibility (Sadler, Mannucci et al. 2000) and should not form part of a standard diagnostic workup for the disorder.

1.13.1.2 Closure time (CT)

CT is the determination of VWF dependant platelet function using the PFA-100TM (Platelet Function Analyser). This involves a high sheer stress system and can produce sensitive and reproducible results to screen for VWD, however the CT is normal in type 2N VWD (Sadler, Mannucci et al. 2000). Studies have shown that the PFA-100 has a higher sensitivity than the BT for detecting moderately severe VWD, however has reduced sensitivity for identifying cases of mild type 1 VWD (Cattaneo, Federici et al. 1999; Nitu-Whalley, Lee et al. 2003; Posan, McBane et al. 2003).

1.13.1.3 Platelet count

The platelet count, which is part of a full blood count, is usually included in the initial screening tests for VWD. It is not directly affected by other types of VWD but may be decreased in VWD type 2B, especially in relation to surgery or pregnancy (Sadler, Mannucci et al. 2000).

1.13.1.4 Activated Partial Thromboplastin Time (APTT)

The APTT may be found to be prolonged in patients with VWD as plasma levels of FVIII commonly depend on the VWF levels, therefore it is useful in identifying type 3 and type 2N where FVIII is always reduced. However, it is not sensitive for screening milder forms of VWD where FVIII levels are normal or near-normal (Sadler, Mannucci et al. 2000).

1.13.2 Specific Tests

Plasma VWF and FVIII levels can be influenced by many physiological parameters such as age, ethnic group and hormonal status. The observed normal values show a wide range from 40 to 240 IU/dL and diagnosis may not always be clear-cut as one or more of the activities of FVIII/VWF may be

borderline normal or even normal. It is often necessary to repeat testing on at least three occasions, as the concentration of both proteins can be affected by a number of endogenous factors (e.g. intercurrent illness). The ABO blood group is also related to both FVIII and VWF levels. Those with blood group O have FVIII/VWF levels that are 15-20% lower than those with blood group A; this should be taken into account when investigating people with borderline results (Gill, Endres-Brooks et al. 1987). In some patients with normal or borderline VWF levels, a primary platelet disorder should also be taken into account, as the clinical presentation of bleeding can be identical.

1.13.2.1 Factor VIII Activity (FVIII:C)

Factor VIII secretion and plasma half-life are dependant on VWF because of its function as a carrier and stabiliser of FVIII in the plasma. Therefore FVIII:C levels usually mirror those of VWF:Ag. Low FVIII levels relative to VWF levels (FVIII/VWF:Ag <1) suggests type 2N disease, which is characterised by poor interaction between VWF and FVIII.

FVIII:C is measured by either a one-stage or two-stage clotting assay based on the APTT method or by a chromogenic assay based on the generation of activated FX (Sadler, Mannucci et al. 2000).

1.13.2.2 von Willebrand Factor Antigen (VWF:AG)

Plasma VWF antigen (VWF:Ag) measures the level of VWF in circulation using immunological methods. Previous electrophoretic immunoassays have been replaced with more sensitive enzyme immunoassays. Decreased levels of plasma VWF:Ag are seen in patients with quantitative VWF defects but may be normal in type 2 variants (Sadler, Mannucci et al. 2000). Type 3 VWD is characterised by very low levels of VWF:Ag (<1 IU/dL) as this indicates an almost total absence of VWF in the plasma.

1.13.2.3 Ristocetin CoFactor Activity (VWF:RCO)

The VWF:RCO assay is used to determine VWF activity by measuring the interaction between VWF and platelet GPIb in the presence of the antibiotic ristocetin. Aggregometry is used to assess platelet agglutination following

addition of ristocetin to a suspension of normal platelets and patient plasma (Macfarlane, Stibbe et al. 1975). The rate and extent of platelet agglutination is dependant upon the concentration of VWF. This is a useful VWD screening test; however it is limited by poor reproducibility.

Quantitative type 1 and type 3 VWD and VWF variants with conserved platelet-dependant function (type 2N) show a normal ratio of VWF:RCo/VWF:Ag. Although type 1 disease can also be identified by an equal reduction in both the activity and amount of VWF (VWF:RCo/VWF:Ag ratio of >0.6). Some type 2 forms of VWD are indicated by a disproportionate reduction in VWF activity compared to VWF levels (VWF:RCo/VWF:Ag ratio of <0.6) as normal ristocetin-mediated platelet agglutination is dependant upon the largest multimeric forms of VWF.

1.13.3 Discriminating Tests

Following a diagnosis of VWD, these tests are required to identify a specific subtype.

1.13.3.1 Ristocetin-Induced Platelet Aggregation (RIPA) Assay

The primary haemostatic function of VWF is best assessed by its ability to support ristocetin-induced platelet aggregation in vitro. This involves measuring the ability of a patient's VWF to bind to platelets in vitro in platelet-rich plasma at varying concentrations of ristocetin. RIPA is undetectable with any ristocetin concentration in VWD type 3, but may be normal in patients with VWD type 1. Type 2B disease can be characterised by RIPA as they have VWF that has a higher affinity for GpIb α receptor binding and so can stimulate platelet aggregation at lower concentrations of ristocetin (<0.8 mg/ml) compared to other patients with VWD (>1.2 mg/ml).

1.13.3.2 VWF Collagen Binding (VWF:CB)

The VWF:CB assay measures the binding of VWF to collagen and measures preferentially the HMW multimers of VWF and complements the results of the VWF:RCo. VWF:CB activity is measured using ELISA plates coated with collagen and the amount of patient VWF that binds is quantified and compared

to a plasma standard. This assay is helpful in distinguishing certain subtypes of type 2 VWD from type 1 as it has the ability to preferentially bind HMW multimers which are lacking in type 2A and type 2B VWD (Favaloro 2000).

1.13.3.3 VWF Multimer Analysis

VWF multimeric analysis is used to visualise the pattern of VWF multimers in plasma in order to distinguish VWD subtypes. VWF multimer analysis can be performed by sodium dodecyl sulphate-agarose discontinuous gel electrophoresis with enzyme-labelled antibodies. The VWF multimers are visualised using enhanced chemiluminescence kits followed by autoradiography (Cumming and Wensley 1993). Normal plasma VWF is visible as a series of bands separated in size by the mass of two VWF subunits and range in size from 500 to >10,000 kDa.

Type 1 patients exhibit the full range of VWF multimers but they may be reduced in intensity whereas they are absent in patients with type 3. Type 2A and 2M both have low VWF activity compared to the amount of VWF present therefore they can be distinguished from one another by assessing the size distribution of multimers, as type 2A VWD patients lack HMW VWF multimers whereas type 2M do not.

1.13.3.4 VWF-FVIII Binding Assay

This assay measures the affinity of FVIII-VWF binding (VWF:FVIIIIB) by isolating patient VWF by immunoadsorption, with subsequent binding of purified, exogenous, FVIII. The bound FVIII is then measured either by a chromogenic assay or an ELISA method. This test can be used to confirm a diagnosis of VWD type 2N in those with a low FVIII/VWF:Ag from mild or moderate forms of haemophilia A.

1.13.4 Treatment of VWD

Treatment for VWD patients is targeted at correcting the functional deficiency of VWF and increasing low FVIII levels (Favaloro, Lillicrap et al. 2004). The two main treatment options are DDAVP and transfusion with blood products that contain the FVIII-VWF complex (replacement therapy). DDAVP is most useful

in treating patients with type 1 VWD however patients with type 3 VWD are usually unresponsive. Unresponsive patients are generally treated with replacement therapy.

Other non-transfusional treatments include antifibrinolytic drugs such as tranexamic acid and hormonal therapy (i.e. oral contraceptive). Antifibrinolytic drugs can either be given alone or in combination with other therapies and are effective in the treatment of nasopharyngeal, oral or gastrointestinal bleeding (Mannucci 1998).

1.14 MOLECULAR PATHOLOGY OF VWD

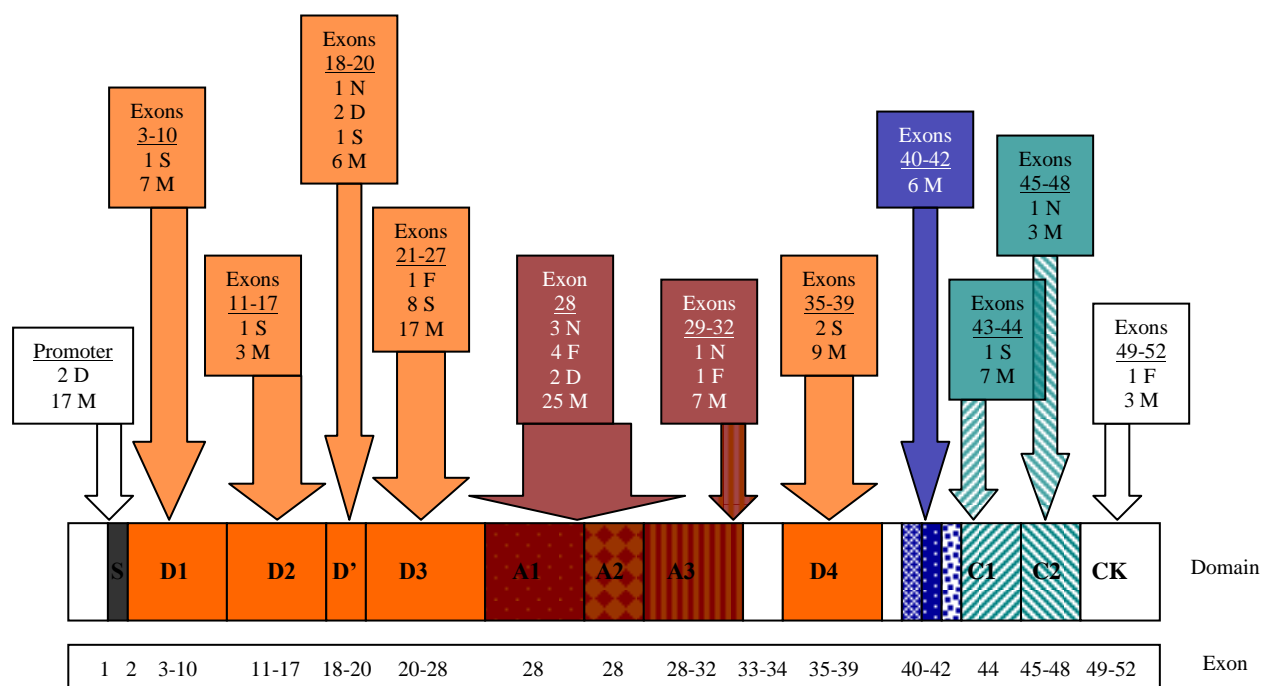
Mutations identified in *VWF* are generally thought to be the main cause of VWD as they have been found in coding regions which alter the expression, processing, structure and function of VWF and mature VWF (Sadler 1994). However, it is now understood that VWD is not only associated with mutations in *VWF*. The underlying genetic defects associated with types 2 and 3 VWD are well defined but the molecular pathogenesis behind type 1 VWD still remains inconclusive, particularly in milder cases. There have been many mutations reported to be associated with the various subtypes of VWD and a database of known *VWF* mutations and polymorphisms is maintained by The International Society on Thrombosis and Haemostasis Scientific and Standardisation Committee (ISTH SSC) (Ginsburg and Sadler 1993). The database is maintained by the School of Medicine and Biological Sciences at the University of Sheffield and can be accessed at <http://www.vwf.group.shef.ac.uk>.

1.14.1 Type 1 VWD

As mentioned earlier the underlying molecular mechanisms of type 1 VWD have only been identified in a small proportion of patients and lacking a common mechanism causing the disease state. Three large population studies in the UK, Canada and Europe have been carried out to investigate the genetic defects causing type 1 VWD and reported inconclusive results (Cumming, Grundy et al. 2006; Goodeve 2007; James, Notley et al. 2007). Therefore supporting the hypothesis that the type 1 VWD phenotype does not always

correlate with the genotype rendering the genetic diagnosis of type 1 as limited (Collins, Cumming et al. 2008). According to the ISTH *VWF* mutation database the majority of mutations or candidate mutations associated with type 1 VWD are missense mutations. The rest are made up of small deletions, splice site mutations, promoter region mutation, insertions, duplications and nonsense mutations. These reported mutations are spread throughout the entire length of the gene and not restricted to a specific region of *VWF* (see Figure 1.7).

Figure 1.7: Diagram representing the distribution of type 1 VWD mutations in *VWF* (Cumming et al. 2010), used with the authors' permission. N = nonsense, F = frameshift, D = small deletion, S = splice site, M = missense



The three recent population studies have highlighted the need for further studies to be carried out to examine the cause of type 1 VWD which include the possibility of small or large deletions (see section 1.17), or splicing variants not caused by changes at the invariant splice sites. Although it was also made clear that there may be no apparent mutational mechanism responsible for a large proportion of type 1 VWD.

1.14.2 Type 2 VWD

The majority of the mutations are found within the VWF domain A2 which is located in exon 28 of *VWF*, with fewer in VWF domain A1 (Schneppenheim, Budde et al. 2001) (see Figure 1.8). The majority of mutations reported on the ISTH *VWF* database are missense, with a smaller percentage caused by frame shifts, insertions and deletions. The most commonly reported missense mutations are c.Arg1597Trp and c.Ser1506Leu within the A2 domain which account for up to a third of type 2A cases (Meyer, Fressinaud et al. 1997; Nichols and Ginsburg 1997).

1.14.2.1 Type 2A VWD

Two classes of mutations are known to produce type 2A VWD, referred to as group 1 and group 2 mutations (Lyons, Bruck et al. 1992; O'Brien, Sutherland et al. 2005). Group 1 mutations cause defective intracellular transport of VWF and damages the assembly, storage and secretion of large VWF multimers in both the plasma and platelet compartments. Group 2 mutations do not affect VWF assembly or secretion but make the multimers more sensitive to proteolysis in plasma (Dent, Berkowitz et al. 1990). In both cases of type 2A VWD mutations, the mutant subunit results in the characteristic loss of large VWF multimers.

1.14.2.2 Type 2B VWD

All of the mutations associated with type 2B VWD are located in the A1 domain of VWF which contains the Gplb binding site. Nearly all of these are missense mutations with the most frequently reported mutations being p.Arg1306Trp, p.Arg1308Cys, p.Val1316Met and p.Arg1341Gln accounting for 70% of type 2B VWD (Cooney, Nichols et al. 1991; Lillicrap, Murray et al. 1991; Eikenboom, Reitsma et al. 1994; Casana, Martinez et al. 1998).

1.14.2.3 Type 2M VWD

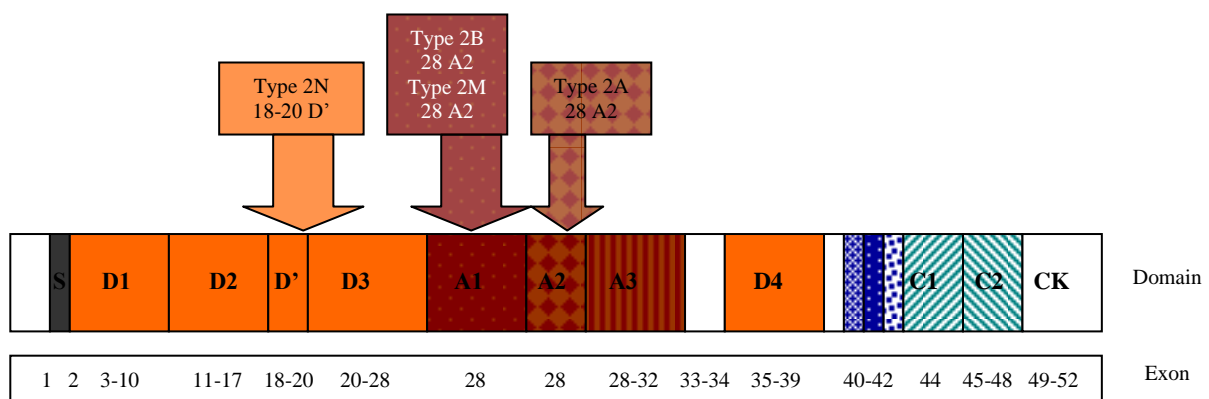
The majority of mutations associated with type 2M are located in the A1 domain which mediates VWF interaction with platelet Gplb. Therefore these mutations, which are primarily missense, result in impaired binding of VWF to platelet Gplb (Meyer, Fressinaud et al. 1997; Casana, Martinez et al. 1998; Casonato,

Pontara et al. 2006). Another variant known as VWD type 2M (Vicenza) displays extra large VWF multimers in plasma and low VWF antigen and activity, caused by candidate mutations in exon 28 of *VWF* (Casonato, Pontara et al. 2006).

1.14.2.4 Type 2N VWD

Individuals with type 2N VWD are expected to be homozygous, compound heterozygous or pseudo-homozygous (i.e. inheritance of one allele containing a type 2N VWD mutation and one allele with a *VWF* null mutation). According to the ISTH database the majority of mutations associated with type 2N are missense and are located in exons 18-20 of *VWF*. The most commonly reported mutation is p.Arg854Gln and has been shown to interfere with VWF binding to FVIII which causes low level FVIII levels (Hilbert, Jorieux et al. 2004). There have been reports of mutations associated with type 2N which lie outside the FVIII binding domain, however these may cause key structural changes to VWF which result in loss of ability to bind to FVIII effectively (Gu, Jorieux et al. 1997; Casonato, Sartorello et al. 2003). Co-inheritance of a type 2N VWD mutation with a type 1 VWD allele has also been reported which may contribute to the variable expressivity of type 1 VWD (Eikenboom, Reitsma et al. 1993; Casonato, Gaucher et al. 1998).

Figure 1.8: Diagram representing the distribution of type 2 VWD mutations in *VWF* (Cumming et al. 2010) used with the authors permission.



1.14.3 Type 3 VWD

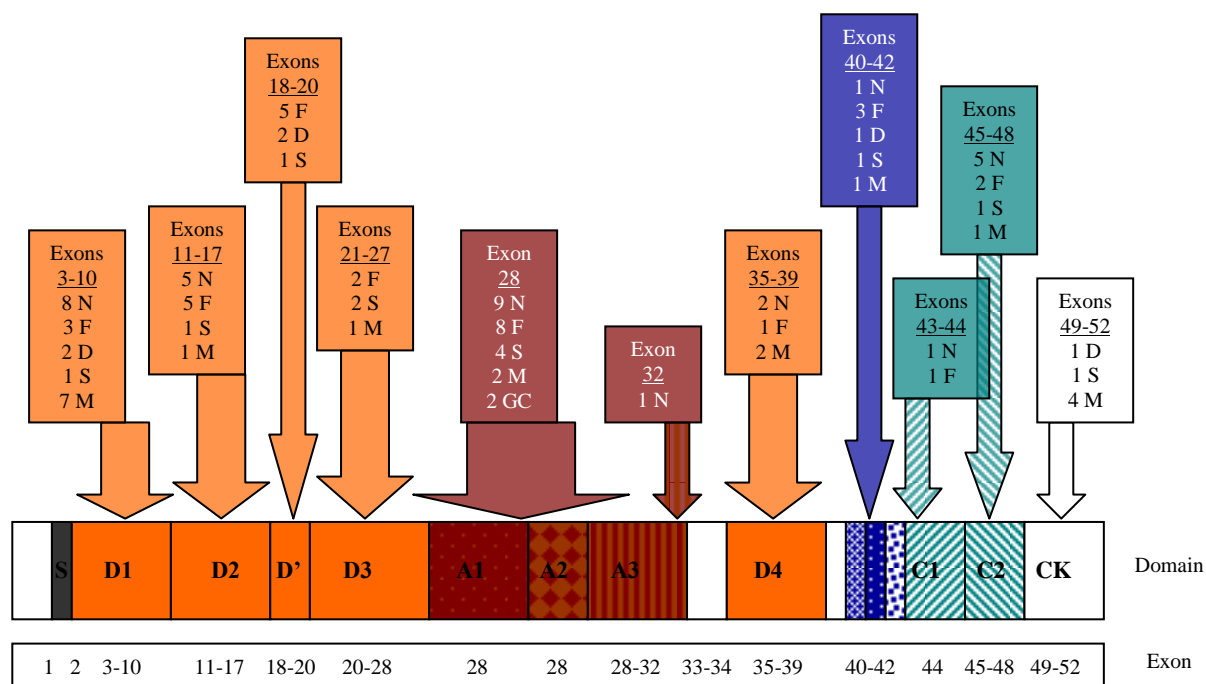
Type 3 VWD arises because of homozygous or compound heterozygous recessive mutations at the *VWF* locus that produce null alleles. A wide spectrum of mutations are associated with type 3 VWD and can be found throughout the 178 kb length of *VWF*, including the promoter region, coding and non-coding regions, and the 5' and 3' untranslated regions (see Figure 1.9)

Initially type 3 VWD was presumed to arise from large deletions of the gene encoding *VWF*, however according to the ISTH *VWF* database only 12 of 109 type 3 VWD mutations are large gene deletions. Large deletions are generally defined as ranging from the deletion of one or more exons of *VWF* and have been shown to be a relatively infrequent cause of type 3 VWD (Baronciani, Cozzi et al. 2000; Baronciani, Cozzi et al. 2003). However a recent study by Sutherland, Cumming et al. (2009) identified a large deletion of *VWF* exons 4-5 in a cohort of type 3 VWD patients from the north-west of England. This mutation was seen in 7 of 12 Caucasian type 3 VWD patients from unrelated families but was absent in nine patients of Asian origin with haplotype analysis indicating a possible founder origin of the mutation in families originating from Britain. This mutation was found to be *Alu*-mediated which has also been seen in at least two other deletions in type 3 VWD as *Alu* repeats are known to promote unequal homologous recombination which results in deletion mutations (Xie, Wang et al. 2006; Mohl, Marschalek et al. 2008). Heterozygous carriers of *VWF* deletions are usually found to be clinically asymptomatic and have normal levels of VWF which proves that the presence of a single *VWF* locus can supply sufficient functional VWF. Although it has been reported that some carriers do have reduced levels of VWF:Ag and VWF:RCO and that some heterozygous carriers of type 3 VWD demonstrate a mild type 1 VWD phenotype (Peake, Liddell et al. 1990; Anvret, Blomback et al. 1992).

The exon 4-5 deletion discussed above has also been found in a number of type 1 VWD families segregating with VWD in an autosomal dominant fashion and exerting a dominant-negative effect to produce a type 1 VWD phenotype. The most common mutations are total or partial *VWF* deletions, nonsense, splicing and frame shift mutations. Deletion mutations result in the elimination

of vital sequences of *VWF* and range from the deletion of one or more exons of *VWF* to the deletion of the entire gene (see section 1.17). The presence of large or complete *VWF* deletions or other gene deletions which lead to the production of a truncated or absent protein, can cause the development of alloantibodies to VWF following the transfusion of VWF replacement therapy products. These anti-VWF alloantibodies can produce severe anaphylactic reactions in patients who are receiving replacement treatment at the time (Mannucci, Tamaro et al. 1987). Therefore it is important to identify the underlying VWF defect in type 3 VWD patients to aid the assessment of risk for inhibitor development.

Figure 1.9: The distribution of type 3 VWD mutations in *VWF* (Cumming et al. 2010) used with the authors permission. N = nonsense, F = frameshift, D = deletion (small or whole exon), S = splice site, M = missense, GC = gene conversion event.



Nonsense mutations or small deletions or insertions associated with a null allele make up approximately 80% of reported *VWF* mutations in type 3 VWD. These types of mutations produce mutant alleles which fail to generate any mature VWF protein leading to a complete quantitative deficiency. A few

studies have identified splice-site mutations associated with type 3 VWD which result in exon skipping in affected VWF mRNA transcripts (Eikenboom, Castaman et al. 1998; Baronciani, Cozzi et al. 2000; Hollestelle, Thinnes et al. 2001; Baronciani, Cozzi et al. 2003). Homozygous or compound heterozygous missense mutations are responsible for 15-20% of type 3 sequence variants (Cumming et al. 2010). Missense mutations specifically located at a cysteine amino acids are expected to have a detrimental effect on mature VWF because cysteine residues aid the formation of disulphide bonds, which are required for the proper function and tertiary structure of the mature VWF protein. Unpaired cysteine residues are recognised to inhibit the intracellular transport of many proteins. It has been shown in expression studies of normal and mutant *VWF* that these mutations can interfere with dimerisation of VWF monomers (Schneppenheim, Krey et al. 1994; Eikenboom, Castaman et al. 1998; Castaman, Eikenboom et al. 2000; Schneppenheim, Budde et al. 2001; Tjernberg, Vos et al. 2004) and impair VWF multimerisation and secretion through a dominant-negative mechanism (Zhang, Blomback et al. 1994; Allen, Abuzenadah et al. 2000; Baronciani, Cozzi et al. 2000; Baronciani, Cozzi et al. 2003; Baronciani, Federici et al. 2008). It is also thought that mutations which cause frame shifts result in defective protein folding hence interfering with normal propeptide processing (Mohlke, Nichols et al. 1996).

Gene conversions between *VWF* and *VWFP* have been identified as making a significant contribution to the molecular mechanism behind type 3 VWD. The conversion event occurs due to the high degree of sequence homology between *VWF* and *VWFP* (97%) and is generally constrained to the region of the 3' part of intron 27 and 5' part of exon 28 of the gene and pseudogene. This is a result of chi sequences (CCTGGTGG) or chi-like sequences (GCTGGTGG) present in the region which encourage DNA recombination (Kenter and Birshtein 1981). The identification of multiple base substitutions within the region of *VWF* corresponding to the *VWFP* sequence signifies that a gene conversion has taken place. A common mutation that has often been seen in association with a gene conversion event in type 3 VWD is the nonsense mutation p.Gln1311X in exon 28 of *VWF* (Casana, Martinez et al. 2000; Surdhar, Enayat et al. 2001; Baronciani, Cozzi et al. 2003; Gupta,

Adamtziki et al. 2005; Gupta, Saxena et al. 2008). The phenotypic expression can differ depending on the length of the gene conversion or whether it is homozygous, heterozygous or compound heterozygous.

1.15 GENETIC TESTING IN VWD

Section 1.13 describes the phenotypic diagnosis of VWD based on the patient's clinical and familial features. The practice guidelines for the molecular analysis of VWD describe how molecular analysis of *VWF* can help to identify the underlying genetic defect responsible for the phenotypic characteristics and help to confirm a diagnosis (Keeney, Bowen et al. 2008). Now that gene sequencing technology is more straightforward and is more widely available it is possible to screen the whole or relevant parts of *VWF* for specific mutations to aid VWD classification. This may be significant when confirming a diagnosis of type 2A, 2B, 2M or 2N, however in cases where the subtype is clear from the phenotype, its clinical use is limited. In type 3 VWD, genetic investigation is particularly beneficial for family studies and prenatal diagnosis. It can also provide important information for clinicians involved in the management of the birth of an affected child and post natal care of the neonate. Genetic counselling is an essential requirement for patients and families with VWD, before and after genetic investigations are undertaken. This is so that individuals understand the inheritance of VWD in their families and the results of subsequent genetic tests (Keeney, Bowen et al. 2008).

In the UK, genetic testing for VWD should be performed in a member laboratory of the UKHCDO-GLN which is a consortium of laboratories who work to peer-defined, peer reviewed standards (Laffan, Brown et al. 2004) to provide a comparable quality of testing nationally.

1.15.1 Diagnostic Strategy

Mutation analysis is carried out by DNA sequencing of the essential regions of *VWF* which is considered to be the gold standard. The essential regions of *VWF* are characterised by the promoter region, exons 1 to 52 together with the splice junctions and flanking sequences, and the 3' UTR region.

Genetic testing and family studies in type 3 VWD are important because of the severity of the bleeding tendency and the autosomal recessive pattern of inheritance. It is particularly valuable for the identification of asymptomatic carriers of type 3 VWD as this cannot be diagnosed phenotypically or by pedigree analysis. If no causative mutations are found in a type 3 VWD patient, or if a patient appears to be homozygous for a given mutation, the presence of an insertion, rearrangement or whole or partial gene deletion on the other *VWF* allele should be taken into account. Southern blot analysis can be used to investigate this, however this is no longer a mainstream routine diagnostic technique (see section 1.9 for other described methods).

For the majority of type 2 VWD cases a diagnosis is made following phenotypic testing, however, now that the molecular basis for type 2 is more clearly understood, genetic testing may help to confirm a phenotypic diagnosis.

1.16 DETECTION OF LARGE SCALE HETEROZYGOUS DELETIONS

CNV's involving large deletions within *VWF* have been reported, although in the past only homozygous deletions have been found in type 3 VWD, as heterozygous deletions (or duplications) are not easily detected during PCR amplification due to the presence of one normal *VWF* allele. The presence of a large homozygous *VWF* deletion is generally predicted to disrupt protein translation and prevent VWF expression. A few reports have identified large heterozygous deletions in types 1 and 2 VWD using linkage analysis which has highlighted unusual polymorphic marker inheritance (Bernardi, Marchetti et al. 1990; Johansson, Hallden et al. 2011) or by using a mutation specific PCR assay for a previously characterised type 3 VWD deletion (Sutherland, Cumming et al. 2009). These mutations tend to be in-frame and function in a dominant-negative manner.

Since standard diagnostic methods, including DNA sequencing, are unable to detect large heterozygous deletions, various techniques can be used to identify mutation status in heterozygous carriers. These include linkage analysis (which may reveal loss of heterozygosity for markers in the deleted region), analysis of platelet derived VWF mRNA (which may reveal an aberrantly sized RT-PCR

product), gene dosage analysis, such as MLPA and Gap or Long PCR protocols (to partially or fully map deletion boundaries) (Keeney, Bowen et al. 2008). MLPA can be used to detect homozygous and heterozygous CNV's and recently has been used to identify additional deletions contributing to VWD in both heterozygous type 2 and in compound heterozygous type 3 patients (Yadegari, Driesen et al. 2011).

1.17 MOLECULAR STUDIES OF LARGE DELETIONS IN VWD USING MLPA

Large deletions are regarded as a rarity for being the underlying genetic defect in VWD, which generally result in a complete absence of VWF protein (Schneppenheim, Castaman et al. 2007). In type 3 VWD, 12 large deletions have so far been reported, ranging in size from a single exon to the entire gene. It is also worth noting that a large deletion spanning all VWF A domains has been identified in type 2A VWD (www.vwf.group.shef.ac.uk).

As discussed above, the identification of large deletions is harder in autosomal conditions, where the presence of a normal allele masks the failure to amplify the mutant one. Therefore false-negative results are often found in carriers of heterozygous large deletions. MLPA as described in section 1.9.1 can be used to successfully identify heterozygous deletions. To date there have been two reports of large deletions identified in patients without a causative mutation using MLPA.

The first report was from Cabrera, Casana et al. (2011) who were investigating the underlying genetic defects of two severe type 3 VWD patients. In the two patients with VWD they identified a large homozygous deletion after the PCR technique failed to amplify from exon 15 to exon 43. However in the parents, sister and a non carrier relative of the patients, all exons successfully amplified. Further analysis using MLPA showed that the patients had an almost total reduction in peaks corresponding to exons 16-43 whereas the area of these same peaks was around 50% in their relatives indicating their heterozygous state. The fact that exon 15 had failed to amplify using PCR indicated that the breakpoint of the deletion was located behind the exon 15 reverse primer used for PCR. Therefore long and accurate (LA) PCR was used to determine the

deletion breakpoints and this showed that the deletion covered from the 5' end of intron 15-3266 bp within intron 43. This gave an explanation for the discordant results between PCR and MLPA as the intron 15 breakpoint was located within the nucleotide sequence annealing to the reverse primer. In conclusion they reported a type 3 VWD family whose patients are homozygous for a large 84 kb deletion and all carriers share the same deletion in the heterozygous state. This has allowed them to offer tailored genetic counselling to all members of the family.

Yadegari, Driesen et al. (2011) performed a study on a cohort of patients in Germany using MLPA analysis to look for large deletions in patients where analysis had failed to identify the causative mutations for their VWD. MLPA analysis was carried out using a kit supplied by MRC-Holland. In total they detected 6 large deletions (5 novel) in their cohort of 23 VWD patients. They found five large gene alterations in patients with type 3 VWD and one in a patient with type 2 VWD, however they did not detect any large deletions in 11 index patients with type 1 VWD in whom the VWF antigen level was more than 33 IU/dL. Therefore they concluded that the presence of large deletions should be taken into account for the diagnosis of type 3 VWD, especially in patients who do not have two clear causative mutations. If sequencing results reveal the presence of homozygous polymorphisms or a homozygous mutation then the possibility of a large deletion should always be taken into consideration. Also, patients diagnosed with type 2 VWD without a *VWF* mutation should also be analysed for large gene deletions and although no large deletions were detected in type 1 VWD, they should be investigated for the presence of possible heterozygous large deletions where no causative mutation has been identified.

These studies show that MLPA can be successfully used to determine the dosage of the 52 exons of *VWF* and can aid the identification of possible carriers with heterozygous deletions/insertions.

1.18 PROJECT AIMS AND OBJECTIVES

Knowledge of the molecular basis of inherited bleeding disorders is required in many cases to enable genetic counselling within families, carrier diagnosis and, in some cases, prenatal diagnosis. Genetic information may also be useful for differential diagnosis, and the nature of the genetic defect can inform clinical management in certain circumstances. The current gold standard for genetic diagnosis in patients with VWD and haemophilia A is direct sequencing of genomic DNA to identify gene mutations. However there are cases where large heterozygous gene deletions or duplications in patients with VWD and in haemophilia A carriers are undetectable by DNA sequencing (Sutherland, Cumming et al. 2009). Therefore the genetic basis of disease in a proportion of cases remains unresolved. Recent studies have shown that the MLPA technique which is based on gene dosage can be applied to detect large gene deletions in haemophilia and VWD and therefore establish carrier status (Rost, Loffler et al. 2008; Lannoy, Abinet et al. 2009; Zimmermann, Oldenburg et al. 2010; Cabrera, Casana et al. 2011; Rafati, Ravanbod et al. 2011; Theophilus, Baugh et al. 2011; Yadegari, Driesen et al. 2011).

The aim of this study was to determine the molecular basis of disease in patients with haemophilia A and VWD using MLPA, where direct sequencing of *F8* and *VWF* failed to reveal causative mutations, and to assess its application in the genetic diagnosis of these disorders.

Objectives:

1. To develop, optimise and validate MLPA assays for the detection of large deletions/duplications in *F8* and *VWF*.
2. To apply these methods to investigate the genetic basis of disease in selected patients with Haemophilia A and VWD.
3. To introduce MLPA as part of the routine diagnostic strategy for Haemophilia A and VWD cases analysed within the Haematology Molecular Diagnostics Centre (MDC) at Manchester Royal Infirmary (MRI).

1.19 PREVIOUS STUDIES RELATED TO THIS PROJECT

1.19.1 General Mutation Detection Strategies for Haemophilia A

The diagnostic strategy used for mutation detection is determined by the severity of the haemophilia in the index case. The initial approach for severe haemophiliacs is to screen them for the intron 22 inversion mutation, followed by the intron 1 inversion mutation, which would identify the underlying genetic defect in 45-50% of severe haemophilia A patients. Following this, full mutation analysis of the essential regions of *F8* is carried out on the remainder of severe haemophilia A cases. The strategy for mild and moderate haemophilia cases is that full sequence analysis is performed as there is no common mutational mechanism linked to these forms. This general diagnostic strategy had been applied previously to the patients included in this study and the methods used are outlined below.

1.19.1.1 Intron 22 inversion detection by Long PCR

A Long range PCR method was used for the detection of the *F8* intron 22 inversion. The method has been modified from standard long range PCR methods to include the addition of DMSO (Dimethyl sulfoxide) to enable read through of a high GC content upstream of *F8*. The ABgene Extensor PCR kit is also used to simplify the reaction set-up procedure from the original publications of the method. The multiplex PCR protocol relies on overlapping PCR to generate a constant 10 kb PCR product which appears with all templates, and acts as a control to show that the reaction has worked efficiently. The large fragments produced by long-range PCR are analysed on conventional agarose gels (0.7% w/v for the intron 22 PCR). The largest amplification product seen using this method is 12 kb. Full method details, including full primer sequences, are available in the paper by (Liu, Nozari et al. 1998).

1.19.1.2 PCR Amplification of the F8 Intron 1 Inversion

A dual PCR assay was used to detect the intron 1 inversion. Both are standard PCR assays designed to amplify independently sequence flanking the int1h regions (from intron 1 homologous sites). *Int1h-1* specifies the assay for the copy in *F8* and *Int1h-2* the homologous region 140 kb more telomeric.

Full details for this method are available in the paper by Bagnall, Waseem et al. (2002).

1.19.1.3 Direct DNA Sequencing

Full DNA sequence analysis of the coding regions of *F8* was used to identify unknown Haemophilia A mutations in index cases and their relatives. The protocol utilised custom-designed PCR primers that allows rapid DNA identification and sequencing of *F8*. Thirty-seven sets of primers have been designed to amplify the essential regions of *F8* under the same thermal cycling conditions. Both forward and reverse primers for each PCR have a common tail (N13F and N13R respectively) to facilitate the use of one complementary sequencing primer in all forward and reverse strand sequencing reactions. This was an in-house protocol developed within the MDC.

1.19.1.4 Detection of Partial Gene Deletions

Large deletion mutations are easily identified in affected males due to failure of amplification of the missing regions of *F8*. However, in carrier females it is especially difficult to identify these mutations as the failure to amplify part of *F8* will be masked by the normal gene sequence on the corresponding allele.

Previous diagnostic approaches developed in the MDC could be applied to this problem. Large deletions were identified in affected males by the absence of amplification products for one or more exons using the standard set of 37 PCR primer pairs. The presence of a partial gene deletion was confirmed by combining the forward and reverse primers from the exons either side of the expected breakpoint in a PCR reaction or long PCR reaction (for method details see Keeney et al. 2007). This enables fragments over 2.5 kb long to amplify across the breakpoint. The ABgene Extensor Hi-Fidelity PCR Master Mix was used for long PCR protocols, however if this method failed to amplify any fragments, the *TaKaRa Ex Taq* 10x PCR Buffer, dNTP (deoxyribonucleoside triphosphates) mixture for PCR (TAKARA BIO INC) was used under the same PCR and thermal cycling conditions.

The size of the fragments produced were analysed on an agarose gel. In affected males with the deletion, a smaller than expected amplified fragment was seen. Long PCR amplification was performed for carrier females and the presence of a deletion was revealed by a larger band representing the normal allele and a second smaller band representing the deleted allele. Normal control samples only produced the larger band representing the normal allele, or it failed to amplify since the product would be too large. In these cases a gap PCR approach was used to confirm the presence of the normal allele.

DNA sequencing was used where possible to confirm the identity of the fragments produced, and to investigate the position of the partial gene deletion breakpoint. Sequence traces from patients with partial gene deletions were compared with the normal *F8* sequence to identify the position of the deletion breakpoint.

1.19.2 General Mutation Detection Strategies for VWD

There are a number of mutations that are recurrent in certain VWD subtypes (in particular in type 2 VWD), therefore patients can be directly screened for relevant characterised mutations or for previously identified familial mutations in *VWF*. The following diagnostic strategies had been applied previously to the type 1 and type 3 VWD patients included in this study and the methods used are outlined below.

1.19.2.1 Type 1 VWD

As part of a UK national type 1 VWD study, all index cases entered were initially screened for mutations in exon 27 and exon 28 of the *VWF*. This was performed to screen for any type 2 VWD or qualitative variant mutations. Exon 28 is the largest exon in *VWF* and has the most described mutations identified in type 1 and type 2 VWD (<http://www.vwf.group.shef.ac.uk/>). The remaining index cases with no mutation identified within these regions were screened for mutations in all 52 exons and the promoter region of *VWF* by DNA sequencing.

1.19.2.2 Type 3 VWD

As part of the UK type 3 VWD study, patients had been screened for mutations in *VWF* by direct DNA sequencing. In patients where heterozygosity for a single *VWF* change or no detectable *VWF* mutation was present, platelet-derived *VWF* mRNA was studied to search for large heterozygous deletions or cryptic splice site mutations that could not be detected using conventional PCR techniques. Platelet-derived *VWF* mRNA studies had also been previously carried out on patients from the UK national type 1 VWD study where no mutation had been identified to account for the bleeding phenotype.

1.19.2.3 Screening of VWF

Fifty one previously designed in-house primer pairs, each consisting of a forward and reverse primer, were used in order to amplify the essential regions of *VWF* including the 5' promoter region, 52 exons, flanking intronic regions and 3' untranslated region. Amplification of genomic DNA was performed by PCR (Saiki, Scharf et al. 1985; Saiki, Gelfand et al. 1988) using standard thermal cycling conditions followed by direct DNA sequencing utilising automated capillary electrophoresis. The DNA sequence data generated were analysed using comparative sequence analysis software and any sequence variants were compared with the published *VWF* sequence (<http://www.vwf.group.shef.ac.uk/>).

Exon 15 of *VWF* had previously shown to be difficult to amplify using conventional PCR techniques as it is extremely GC rich in content, which leads to the formation of secondary structures, such as hairpin loops. Therefore this causes problems when trying to amplify genes using standard PCR approaches. Previous work by Cumming, Grundy et al. (2006) had been carried out to optimise the PCR reaction using the TripleMaster® PCR System (QIAGEN® LongRange PCR System, supplied by Qiagen UK) and the annealing temperature of the primers at 57 °C. The cycling conditions for the PCR reaction were also modified from the standard PCR conditions according to the manufacturer's instructions.

Exon 28 of *VWF* had also been shown to be difficult to amplify due to its large, highly polymorphic nature. It proved difficult to design primers that would successfully amplify appropriate segments of the exon without the primers lying over a polymorphism, which would mean that pathogenic mutations would potentially be missed due to allele dropout. Therefore primers were designed to amplify the whole of exon 28 producing a product that is 1556 bp in size. Consequently, sequence analysis of this product had to be performed in two segments using overlapping internal primer pairs which allowed forward and reverse sequence analysis of the full length of exon 28.

1.19.2.4 Identification of the VWF exon 4 and 5 deletion

Previous work by Sutherland et al. (2009) had led to the identification of a large deletion mutation encompassing exons 4 and 5 of *VWF* at the genomic level in the homozygous state in type 3 VWD patients of Caucasian origin. The exact locations of the deletion breakpoints were determined by primer walking. Amplification of the genomic region from exon 3 to exon 6 (26 kb) was performed using 18 primer sets which produced ~1.5 kb overlapping fragments. The truncated product produced using combinations of these primers was analysed by direct sequencing. This led to the design of further primers to amplify a smaller product (1084 kb) so that sequencing could be performed to determine the precise deletion breakpoints. Development of a multiplex PCR enabled the homozygous/heterozygous status of the deletion to be confirmed, as described in Sutherland et al. (2009). For those patients in whom the exon 4 and 5 deletion mutation was identified, haplotype analysis was performed to investigate the possibility of a founder effect for this mutation.

CHAPTER 2: MATERIALS AND METHODS

All analyses were performed at the MDC in MRI which is a Clinical Pathology Accreditation (CPA) laboratory. The MDC is a participant in the UK national external quality assurance scheme (NEQAS) for haemophilia and VWD diagnostic screening through direct DNA sequencing. The appropriate guidance on ethics had been obtained. No specific ethical approval was required for the project as the research carried out was an extension of the patients' diagnostic analysis and all patients whose specimens were used had given appropriate consent at the time of collection. Statistical analysis was not used in this project as the cohort of patients being studied have had the conventional mutation mechanisms eliminated, and therefore pre-selected on this basis for the investigation of the presence of partial gene deletions. For previous work in support of this study, see Introduction section 1.19.

2.1 SELECTION OF PATIENTS

2.1.1 Haemophilia A

Patients were selected from those who had given informed consent for the genetic diagnosis of haemophilia A referred to the MDC for investigation. Full *F8* sequence analysis has been carried out at the MDC since 2003 on all referred patients. All index cases with suspected partial gene deletions detected by absence of amplification of one or more exons were chosen for further study by MLPA. The patient cohort also included index cases where no causative mutation had been identified by DNA sequence analysis. Carrier female relatives of index patients selected were also included (where samples were available). The selected patient group included those with mild, moderate and severe types of haemophilia A. Nineteen patients were investigated in the study which included hemizygous males who could not be detected by conventional techniques and carrier diagnosis of their relatives (see Results Table 3.3).

2.1.2 Type 1 VWD

Patients were selected individuals who had previously been recruited through the UKHCDO clinical network into a UK national study of type 1 VWD. Each

patient had a historical diagnosis of type 1 VWD made by the referring centre and had previously sent blood samples to the MDC for mutation analysis of *VWF*. In total, sixteen patients were investigated by MLPA which included those with suspected partial gene deletions and those where no causative mutation had been previously identified (see Results Table 3.5). Ethics approval, including permissible extension of the work into which the current study fitted, had previously been obtained for the UK national study from the Multi-centre Research Ethics Committee (MREC) for Wales (reference number: MREC 01/9/32).

2.1.3 Type 3 VWD

Patients were selected individuals who had either previously been invited to participate in a study of type 3 VWD in Manchester or had been referred to the MDC for genetic diagnosis. Each patient had been previously diagnosed with type 3 VWD at MRI and had sent blood samples to the MDC for mutation analysis of *VWF*. All patients with the novel exon 4-5 deletion (c.221-977_532+7059del mutation) discovered in this study were included for additional study by MLPA. In total, ten patients were investigated by MLPA including those previously found to have the exon 4-5 deletion and those with a small deletion, duplication or point mutation (see Results Table 3.4). Ethics approval, including permissible extension of the work into which the current study fitted, had previously been obtained for the study of type 3 VWD patients in Manchester from the Central Manchester Local Research Ethics Committee (LREC) (reference number: 05/Q1407/191).

2.2 PHENOTYPIC DATA

Phenotypic data were obtained for all patients included in the study. For the haemophilia A patients, phenotypic laboratory results were acquired from individual patient records which included FVIII:C levels and inhibitor status. Phenotypic data for the type 1 VWD and type 3 VWD index cases were attained from the previous respective studies.

2.3 SAMPLE COLLECTION AND PROCESSING

Previously extracted genomic DNA was obtained for patients from an archived bank of DNA samples at the MDC which are stored at -80 °C indefinitely. For each patient, DNA had been extracted from 10-20 ml of EDTA (Ethylenediaminetetraacetic acid 7.2 mg in a 4 ml tube) anticoagulated fresh whole blood. DNA extraction was performed by an ammonium acetate method (the procedure is a modification of the method described by SA Miller of the Analytic Genetic Centre Inc, Denver). Additional patient samples in the form of extracted DNA were analysed from external referral centres where deletions had been detected by other means.

2.4 MLPA® TECHNOLOGY

MLPA is a high-throughput method that can be performed to detect quantitative genomic changes such as deletions or insertions. It allows the copy number of up to 50 different genomic DNA sequences to be determined in one multiplex PCR-based reaction. The MLPA reaction produces a mixture of amplification fragments ranging between 100 to 500 nt in length, which can be separated and quantified by capillary electrophoresis. SALSA® MLPA® kits are sold by MRC-Holland (The Netherlands).

The principle of the MLPA technique is based upon amplification of the MLPA probes themselves, not the sample DNA. Each probe is made up of two adjacent oligonucleotides which hybridise to their target sequence in genomic DNA and are then ligated. Amplification only occurs following successful ligation of the probes, which then produce exon-specific PCR products of differing lengths. In comparison to conventional multiplex PCR which requires one pair of primers for each fragment to be amplified, only one pair of PCR primers is used in the MLPA PCR reaction, making it a more robust system. This is because in standard multiplex PCR, primers are present in large amounts in the reaction mixture which can lead to various technical complications. Firstly since the efficiency of different primers usually varies, it is difficult to determine the relative quantitation of target sequences. In addition, small differences in reaction conditions often result in large differences in the results obtained. Following the resulting amplification,

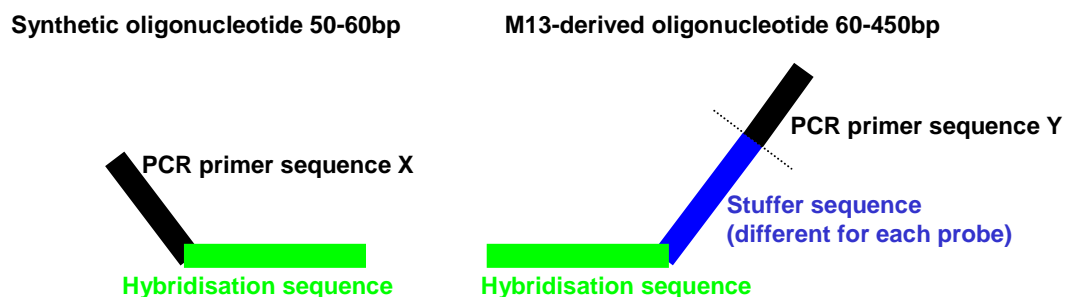
products ranging between 130 and 480 nt in length are separated by capillary electrophoresis. The peak patterns obtained can be compared to reference samples (see 2.6.3.1) to see which sequences show aberrant copy numbers.

2.4.1. MLPA Probe Chemistry

MLPA probes consist of two oligonucleotides, each containing a PCR primer sequence and a sequence complementary to the target, known as the hybridisation sequence (see Figure 2.1).

The two probes hybridise immediately adjacent to each other. When the probes correctly hybridise to the target sequence they are ligated by a thermo stable ligase enzyme. The PCR primers exponentially amplify the ligated probes. One of the primers is labelled with a fluorescent dye to visualise the amplification product of the probe. Probe oligonucleotides that are not ligated only contain one primer sequence therefore they cannot be amplified exponentially and will not generate a signal. The non-ligated probes do not need to be removed therefore the MLPA assay can be analysed in a single tube. Sequencing electrophoresis is used to separate the resulting PCR products. Each MLPA probe set is designed so that the length of each product is unique and can be easily identified when electrophoresed through a gel. The difference in size is achieved with the help of a stuffer sequence.

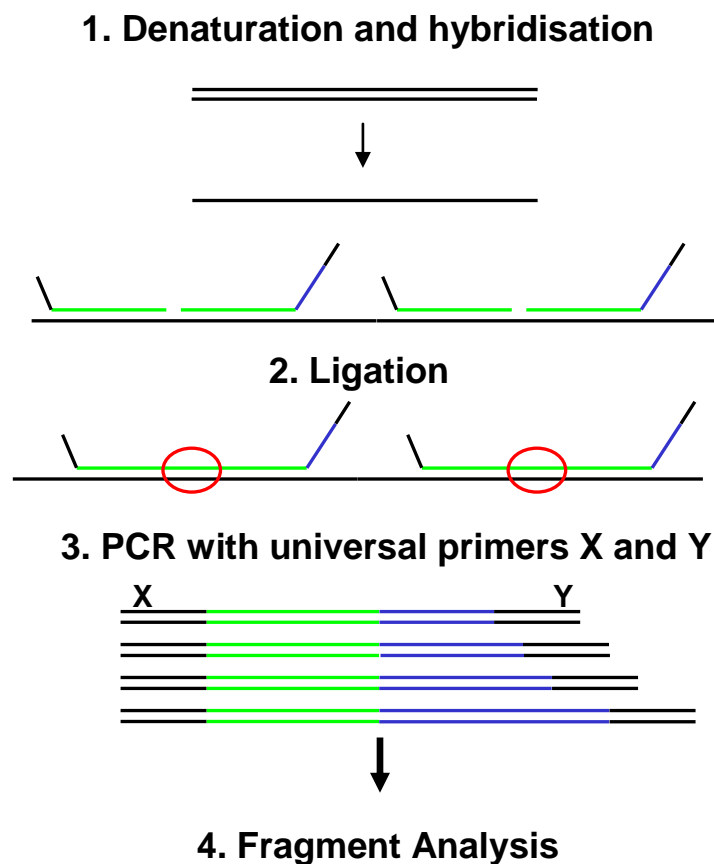
Figure 2.1: MLPA® probe set



2.4.2 MLPA Reaction

The MLPA reaction can be separated in five major steps: 1) DNA denaturation and hybridisation of MLPA probes; 2) Ligation; 3) PCR reaction; 4) Fragment analysis by electrophoresis; and 5) data analysis (see Figure 2.2). During step 1 the DNA sample is denatured to single strands and incubated overnight with the MLPA probe mix. Each MLPA probe consists of two sequence specific oligonucleotides which hybridise to immediately adjacent target sequences. In step 2, the oligonucleotides are ligated with a thermostable ligase but only if they have successfully hybridised to their target sequence. In step 3, a universal primer pair is added to the reaction. The successfully ligated probes are amplified to generate exon-specific products of differing lengths which can be used to determine gene dosage. In step 4, the amplification products are separated using capillary electrophoresis.

Figure 2.2: MLPA reaction overview



2.4.3 Experimental set up

2.4.3.1 Reference samples

Identification of deletions and duplications by MLPA is achieved by analysis of relative changes in probe amplification signals, therefore slight differences between experiments and samples used may affect the MLPA peak pattern. Therefore it was necessary to include reference (control) samples in each MLPA experiment. The reference samples used were DNA samples that had been obtained from healthy individuals. The DNA samples were all derived from peripheral blood and had been extracted using the same method as the test samples, as recommended by MRC-Holland.

At least 5 reference samples from individuals with normal copy numbers of the gene to be analysed were included in each run, together with a 'no DNA control' reaction. The 'no DNA control' reaction was carried out on TE (10 mM Tris-HCl pH 8.2 + 0.1 mM EDTA) and was used to show any possible contamination of dH₂O, TE, MLPA reagents, electrophoresis reagents or capillaries.

Positive controls were not included in the MLPA experiments; it is recommended but not essential.

2.4.3.2 Sample treatment

It is recommended that a total quantity of between 50-250 ng of human genomic DNA in a 5 µL volume is used for each MLPA reaction. MLPA results can be affected by extraction methods, handling and storage of DNA. Although MLPA does not require a unique DNA extraction method, it is essential that the extraction method does not leave a high concentration of contaminants such as salt as it is known to interfere with the technique. Also heparinised samples are not recommended for use as traces of heparin are difficult to remove from DNA preparations and can alter the MLPA reaction.

All DNA samples (including reference samples) used were extracted from peripheral blood in EDTA by the ammonium acetate method. The overall DNA concentration of each sample was measured using a NanoDrop ND-1000

spectrophotometer (supplied by LabTech, UK). Each sample was diluted in TE (10 mM Tris-HCl pH 8.2 + 0.1 mM EDTA) to make a working stock of 40 ng/μL concentration at least 24 hours (h) before use. Some DNA samples which were considered to be of poor quality were diluted in TE to a working stock of 10-20 ng/μL. This was due to the premise that reducing the amount of sample DNA also reduces the amount of impurities and often has a favourable effect on the MLPA result, although too little DNA in the reaction may lead to failure of ligation and amplification. An RNase treatment was used for some samples which consistently failed (see section 2.6.1 for details). This can be used for every sample however it is only essential when screening for a gene that is very highly expressed in the sample tissue under investigation.

2.5 SALSA® MLPA® KITS

The SALSA® MLPA® kits are sold by MRC-Holland. The following kits were used in this study:

1. Haemophilia A – **SALSA® MLPA® KIT P178-B1 *F8***
2. von Willebrand disease (2 part kit) – **SALSA® MLPA® KIT P011-B1 / P012-B1 *VWF***

The SALSA® MLPA® kits consisted of reagents shown in Table 2.1. They were stored at -20 °C and were shielded from light. Prior to use all buffers and probe mix reagent tubes were thawed on ice, briefly vortexed and spun down for a few seconds to dislodge drops that may have adhered to the lid. Enzyme solutions were never vortexed to avoid inactivation. To reduce viscosity, it was recommended to hold the enzyme solution in your hand for 10 seconds (s) and then mix by pipetting up and down. Mastermix containing the enzyme was also mixed by pipetting gently before use. For details of SALSA® MLPA® probemixes for *F8* and *VWF* see Appendices 1, 2a and 2b.

Table 2.1: Components of SALSA® MLPA® kit

SALSA® MLPA® kit component	INGREDIENTS
SALSA MLPA Buffer (yellow cap)	KCl, Tris-HCl, EDTA and PEG-6000. pH 8.5
Probemix (black cap)	Synthetic oligonucleotides, oligonucleotides purified from bacteria, Tris-HCl, EDTA. pH 8.0
Ligase Buffer A (transparent cap)	NAD (bacterial origin). pH 3.5
Ligase Buffer B (white cap)	Tris-HCl, non-ionic detergents, MgCl ₂ . pH 8.5
SALSA Ligase-65 (green cap)	Glycerol, BRIJ (0.05%), EDTA, β-Mercaptoethanol (0.1%), KCl, Tris-HCl. pH 7.5, Ligase-65 enzyme
SALSA PCR Buffer (red cap)*	Tris-HCl, EDTA, KCL, non-ionic detergents. pH 8.5
SALSA PCR Primer Mix (brown cap)**	Synthetic oligonucleotide, one of which is fluorescently labelled with FAM, dNTP's, TRIS-HCl, KCl, EDTA, BRIJ (0.04%). pH 8.0
SALSA Polymerase (orange cap)	Glycerol, BRIJ (0.5%), EDTA, DTT (0.1%), KCl, Tris-HCl, Polymerase enzyme (bacterial origin). pH 7.5
SALSA Enzyme Dilution Buffer (blue cap)*	Tris-HCl, KCl, EDTA, non-ionic detergent. pH 8.5

* **Not required for one-tube protocol, introduced in June 2011**

****New MLPA PCR primer mix introduced for one-tube protocol**

2.6 MLPA REACTION METHOD

The MLPA reaction method is a two day method with the first day requiring an 18 h overnight incubation.

The use of a thermocycler with a heated lid (99-105 °C) was essential (2720 Thermocycler, Applied Biosystems, UK). Before any initial MLPA experiments were set up, a test was performed on the thermocycler for excessive evaporation, which could be due to problems with the heated lid or brand of tubes (the plastic of the tubes is weak at high temperatures and so can deform). This was carried out by incubating 8 µL water overnight at 60 °C in two different brands of tubes. The next day at least 5 µL water was required to still present at the bottom of the tube. The 0.2 ml 8-strip PCR tubes with individually attached dome caps from Starlab (Cat no. I1402-2900) had the least amount of evaporation had and so were used in all MLPA experiments.

2.6.1. DNA Denaturation (Day 1)

On day 1, 5 µL of DNA sample (50-250 ng) was aliquoted into each tube of 0.2 ml strips (or TE for the 'no DNA control reaction') and spun down in a strip microcentrifuge (PeqLab, Perfect Spin Mini). The tubes were placed in the thermocycler and the MLPA thermocycler program started (see Table 2.2). The sample DNA was denatured for 5 min at 98 °C and cooled to 25 °C before being removed from the thermocycler (see Table 2.2).

2.6.1.1 RNase treatment

If using an RNase treatment, 4 µL of DNA sample was aliquoted into each tube of 0.2 ml strips together with 1 µL of RNase treatment. The tubes were placed in the thermocycler and the MLPA thermocycler program started with an extra RNase treatment step for 30 min at 37 °C prior to denaturation (see Table 2.2).

A working stock of the RNase treatment was made by making a 1 in 8 dilution of the stock RNase A Solution (Promega, code A7973; 4 mg/mL solution) in TE buffer to obtain a concentration of 0.5 mg/mL. This was then kept at -20 °C for future use.

2.6.2. Hybridisation reaction (Day 1)

Prior to the hybridisation reaction, a PCR mastermix was prepared containing MLPA buffer and probemix:

Per reaction:

1.5 µL MLPA buffer (yellow cap)

1.5 µL probemix (black cap)

Following DNA denaturation, the samples were removed from the thermocycler and spun down. Using a multi-channel pipette, 3 µL of the hybridisation master mix was added to each sample tube (the thermocycler program is held at 25 °C for 15 min to allow enough time for the hybridisation master mix to be added). The strips were returned to the thermocycler and the program continued with incubation for 1 min at 95 °C and then for 18 h at 60

°C overnight (see Table 2.2). As part of optimisation, different thermal cycler parameters were assessed to determine which would produce the best results (see Results section 3.2.4).

2.6.3. Ligation reaction (Day 2)

On day two, a ligase mastermix was prepared shortly before the end of the 18 h hybridisation reaction using the following reagents:

Per reaction:

25 µL dH₂O

3 µL Ligase Buffer A (transparent cap)

3 µL Ligase Buffer B (white cap)

1 µL Ligase-65 enzyme (green cap)

Following the 18 h at 60 °C, the thermal cycling reaction was cooled to 54 °C and paused (the pause button on the thermocycler allows a 10 min pause, in which time the ligase master mix should be added while the samples are still on the block). Using a multi-channel pipette, 32 µL of ligase master mix was added to the tubes and mixed gently by pipetting up and down, making sure not to introduce any bubbles. The thermocycler program was continued with 15 min incubation at 54 °C (for ligation) followed by 5 min at 98 °C for heat inactivation of the Ligase-65 enzyme. The samples were then held at 4 °C until removed from the thermocycler.

Table 2.2: Thermocycler Program for the MLPA Denaturation, Hybridisation and Ligation Reactions

Step	Temperature	Time	No. of cycles
RNase treatment*	37 °C	30 min	
DNA Denaturation	98 °C 25 °C	5 min Pause	Held for 15 min
Hybridisation reaction	95 °C 60 °C	1 min pause	Held for 18 h
Ligation reaction	54 °C 54 °C 98 °C 4 °C	pause 15 min 5 min pause	

* Only required if using an RNase treatment

2.6.4. PCR reaction (Day 2)

2.6.4.1 One tube protocol

The following instructions are for the one-tube protocol introduced by MRC-Holland in June 2011. This also required the use of a new MLPA PCR primer mix (FAM: lot number F44 onwards). The complete ligation reaction is now used for the PCR and the PCR should be started at room temperature. For the PCR reaction a polymerase master mix was prepared and stored on ice prior to use:

Per reaction:

7.5 µL dH₂O

2 µL SALSA PCR primer mix (brown cap)

0.5 µL SALSA polymerase (orange cap).

At room temperature using a multi-channel pipette, 10 µL of polymerase mix was added to each tube, mixed gently by pipetting up and down and spun down in a strip micro centrifuge. The strips were then placed back onto the thermocycler and the PCR reaction program started (see Table 2.3). Once the PCR reaction had completed, samples were stored at 4 °C for up to one week in a dark box wrapped in aluminium foil to protect the light sensitive fluorescent dyes, until fragment separation by capillary electrophoresis.

Table 2.3: Thermocycler Program for the MLPA PCR Reaction

Step	Temperature	Time	No. of cycles
PCR reaction	60 °C**	pause	} 35 cycles
	95 °C	30 s	
	60 °C	30 s	
	72 °C	60 s	
	72 °C	20 min	
	8 °C	pause	

****Not required for one-tube protocol, introduced in June 2011**

2.6.4.2 Two-tube protocol

Prior to June 2011, the two-tube protocol was used which required an extra PCR buffer and Enzyme dilution buffer (see Table 2.1). A new set of strip tubes were required for the PCR reaction. A PCR buffer mix was prepared:

Per reaction:

26 µL dH₂O

4 µL SALSA PCR buffer (red cap)

Then 30 µL of the buffer mix was added to each tube with a multi-channel pipette. At room temperature, 10 µL of the ligation product was transferred to the corresponding PCR tube.

A separate polymerase master mix was then prepared and stored until required:

Per reaction:

5.5 µL dH₂O

2 µL SALSA PCR primers (brown cap)

2 µL SALSA Enzyme dilution buffer (blue cap)

0.5 µL SALSA polymerase (orange cap)

The thermocycler was preheated to 60 °C and paused whilst the PCR tubes were placed in the thermocycler. At this stage, 10 µL of the polymerase mix was added to each tube using a multi-channel pipette and mixed gently by

pipetting up and down. The thermocycler program was continued immediately (see Table 2.3).

2.6.5 Fragment Separation by Capillary Electrophoresis (CE)

For fragment separation the amount of MLPA PCR reaction, run conditions and fluorescent label required for analysis all depend on the type of capillary electrophoresis instrument used.

The following equipment and reagents were used supplied by Applied Biosystems, UK:

1. Capillary electrophoresis instrument with fragment analysis software:
ABI-3130x/ with Standard Foundation Data Collection Software
2. High quality formamide: **Hi-Di Formamide**
3. Labelled size standard: **GeneScan™ 500 ROX™**
4. Electrophoresis polymer: **POP-7**

Immediately prior to fragment analysis, a master mix was prepared with the following reagents:

Per reaction:

9 µL Hi-Di Formamide

0.3 µL ROX size standard

A 96 well plate was used for set-up and 9.3 µL of the mastermix was added to each sample well for the required number of samples. Using a multi-channel pipette, 1 µL of the PCR reaction was added to each well containing the mastermix. The plate was then sealed with a silicone mat and centrifuged at 170 x g for 30 s to spin down the contents. The plate was placed onto the thermal cycler for denaturation at 94 °C for 5 min. Immediately after denaturation, the plate was placed on ice to prevent the subsequent re-formation of secondary structures in the product.

The plate was then loaded onto the ABI-3130x/ sequencer and the data analysis parameters selected:

1. Application – **Genemapper DF3HOXLI**
2. Results group – **MLPA_RESULTS.DF3HOXLI**
3. Instrument protocol – **FragmentAnalysis_POP7**

A 16 channel 50 cm capillary was used to analyse 16 samples per automatic loading.

2.6.6 Overview of fragment analysis

The fluorescently labelled fragments are detected using the ABI-3130x/ analyser and then interpreted using GeneMapper v.4.0 (Applied Biosystems, UK) analysis software.

During capillary electrophoresis, molecules from the samples are electrophoretically injected into thin, fused-silica capillaries filled with liquid polymer (POP-7, Applied Biosystems, UK). Electrophoresis of all samples begins at the same time when a voltage is applied across all capillaries. The DNA fragments migrate toward the other end of the capillaries through the electrophoresis medium with the shorter moving faster than the longer fragments. As the fragments enter the detection cell, they pass through the path of a laser beam, which causes the dyes attached to the fragments to fluoresce. The dye signals are separated by a diffraction system and a charge-coupled device (CCD) camera detects the fluorescence. Five different coloured fluorescent dyes can be detected in the one sample. One of the dye colours is used for a labelled size standard (ROX 500, Applied Biosystems, UK) that is added to each lane to allow comparison of samples from lane to lane or gel to gel. The analysis software uses the size standard to create a standard curve for each lane and the length of each dye-labelled fragment can then be determined by comparing it with the standard curve for that specific lane. Data were collected and stored as electropherograms prior to analysis.

2.7 ANALYSIS OF DATA

Initial analysis of the fragment analysis data were performed using the GeneMapper analysis software version 4.0 (Applied Biosystems, UK). This was prior to probe ratio calculation to evaluate the quality of the MLPA experiments by inspecting the MLPA peak patterns (electropherograms).

2.7.1 Evaluating data

Raw peak patterns were first analysed using the GeneMapper software. The MRC-Holland probemixes contain MLPA probes for each exon of the gene (some exons may have more one probe), reference probes and internal quality control fragments. The various control fragments (see Table 2. 4) are included to identify common errors in the MLPA reaction. Once the raw data had been size-called and visualised to obtain the genotypes, it was essential to evaluate the MLPA internal quality control fragments.

Table 2.4: Internal Quality Control Fragments

Name	Length (nt)	PURPOSE
92 nt control	92	Ligation-dependant probe, works in same way to other MLPA probes. Acts as a standard for comparison of other control fragments.
Q-fragments	64, 70, 76, 82	Indicators of when the amount of DNA is too low. Signals which are higher than $\frac{1}{3}$ of the 92 nt control fragment = DNA quantity insufficient.
D-fragments	88, 96	Identify denaturation problems. Signals which are lower than 40% of the 92 nt control fragment = DNA denaturation problems
X and Y fragments	100, 105	Controls for male and female samples

The Q-fragments are four oligonucleotides which are present in all SALSA MLPA probemixes. As they contain both MLPA PCR primer sequences in a single molecule, they are able to amplify in the PCR reaction without the need for hybridisation to the target DNA or ligation. When there is enough sample DNA in the extraction, the Q-fragments are outcompeted by the amplicons of the MLPA probes as they are present in such small quantities. Therefore if the Q-fragment peaks are low or invisible, this indicates that adequate sample

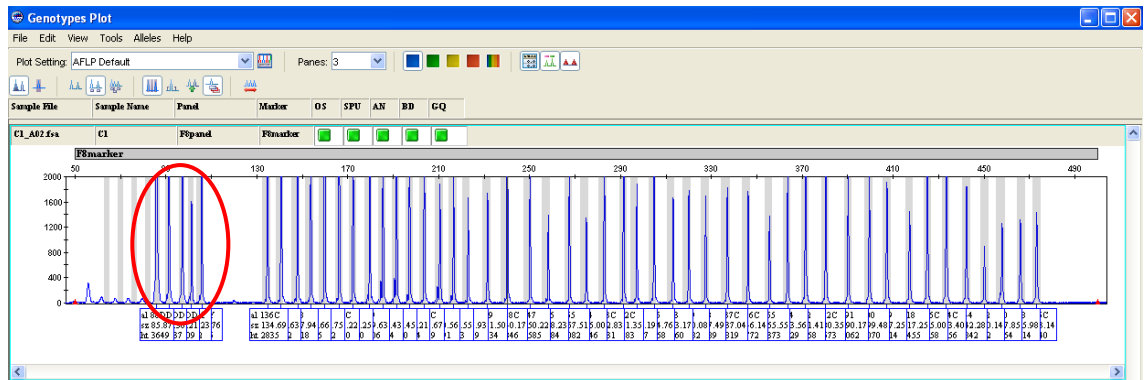
DNA was present and the ligase reaction was successful. However if the peaks of all four Q-fragments are higher than $\frac{1}{3}$ of the height of the 92 nt ligation-dependant control fragment and the MLPA probes, this is an alert that the amount of sample DNA was insufficient or the ligation reaction failed.

The two denaturation D-fragments are synthetic MLPA probes which detect a sequence within a strong CpG island. CpG islands are chromosomal regions which are difficult to denature because of their very high CG content. Incomplete denaturation is signified if the two D-fragment peaks are much lower (<40%) than the 92 nt fragment and the other MLPA probes. This means that results are unreliable for probes that are detecting sequences in or near (<5 kb) CpG islands. This can be as a result of salt contamination of the DNA sample as the melting temperature of DNA is increased at higher salt concentrations. Figure 2.3 illustrates a sample with peaks representing the D-fragments at 88 nt and 96 nt (1st and 3rd peaks, circled in red) which are of equal height to the 92 nt fragment (2nd peak, circled in red) indicating successful denaturation. Exon 1 probes are most often affected by incomplete DNA denaturation as CpG islands are usually located near mRNA transcription start locations. When using the ABI POP-7 polymer for CE, a non-specific fragment of ~87 nt is usually present and may overlap with the 88 nt D-fragment. If this affected analysis of the 88 nt D-fragment then only the 96 nt peak was used.

The X and Y fragments at 100 nt and 105 nt respectively indicate whether the sample is from a male or female. This can be helpful in the case of identifying that a sample swap has occurred if running only male or female samples.

In the no DNA control reaction, only the Q-fragments generate a probe signal. However, MLPA PCR reactions are more prone to having non-specific peaks than a normal PCR reaction as they contain numerous long MLPA probe oligonucleotides, each of which contain a perfect copy of one of the PCR primers. In the majority of cases, the non-specific peaks do not influence the MLPA results as they are outcompeted by the genuine MLPA probe amplification products in the same way as the Q-fragments.

Figure 2.3: Raw peak pattern for a sample tested with the P178-B1 F8 MLPA kit visualised using GeneMapper Software. The probe signals show successful amplification from male DNA indicated by the 4th and 5th peaks (circled in red) representing the X and Y fragments at 100 nt and 105 nt respectively. The 1st, 2nd and 3rd peaks (circled in red) show successful amplification of the fragments for the 88 nt D-fragment, 92 nt control and 96 nt D-fragment respectively.



2.8 INTERPRETATION OF RESULTS

Once the data had passed raw data and peak pattern evaluation, it was ready for data normalisation. The data needs to be normalised as the absolute fluorescence intensities detected by the CE instrument depend on many factors (degradation of the fluorescent dye, variability between capillaries, pipetting accuracy). MLPA normalisation happens in a two step model. First the relative fluorescence intensity of each peak within a sample is determined (intra-normalisation) by dividing the peak area of each probe's amplification product by the total area of only the reference probes in the probemix; in a second step, this relative peak is compared to that in other samples (inter-normalisation) by dividing the intra-normalised probe ratio in a sample by the average intra-normalised probe ratio of all reference samples. The final ratio for a given probe in a test sample compared to the same probe is called the Dosage Quotient of that probe.

MLPA dosage quotients were calculated using spreadsheets designed by the National Genetics Reference Laboratory (Manchester) and were obtained from their website www.ngrl.org.uk/Manchester. The spreadsheets have been

created in Microsoft Excel 2003 and are specific for the different MLPA kits used. They analyse data to produce DQ's in the standard way, however they also include two original features of analysis to help with interpretation.

The first feature is that they are able to generate a likelihood probability of agreeing with one of three hypotheses, that a ligation product is present at either one, two or three copies within the test sample. This figure is a result of comparing the test sample to five normal controls. The amount of variability for each ligation product can be determined from the controls, therefore allowing the probability of deviation from expectation of the test sample to be estimated using the t-statistic.

The second feature measures the standard deviation of the DQ's obtained for all the control ligation products which can be used as a control for the overall quality of the test. Each sample is deemed to be of good quality or poor quality (see Figure 2.4). The samples that have a standard deviation of more than 0.1 are deemed to be poor quality. Retrospective studies by Dr Ruth Charlton, Regional Genetics Service, Leeds (unpublished data) have shown that samples with control standard deviations less than 0.1 shows no overlap between normal, duplicated and deleted ranges. Therefore those samples with higher degrees of variability are excluded which reduces the possibility of making an incorrect diagnosis.

Figure 2.4: Example of the NGRL spreadsheet. This illustrates the appearance of a typical sample where the data quality has been judged to be good (highlighted in green). If the sample quality has been judged to be poor, it will be highlighted in red.

Test 1	Lab No		C1 Xp22.11	CXq27.3
	Operator	C2 Xq27.3	1.00	1.09
	Michelle	C3 Xq13.1	0.92	1.00
	Worksheet	C4 Xq22.1	1.03	1.12
		C5 Xq13.1	0.92	1.00
		C6 Xp21.3	1.03	1.12
	Int QC Stand Dev	C7 Xq28	0.95	1.03
	0.052054129	C8 Xq26.2	0.98	1.07
		C9 Xq22.3	1.02	1.11
		C10 Xp11.3	0.96	1.05
		C11 Xq21.2	1.01	1.10
		C12 Xp21.2	1.04	1.13
		MEAN	0.90	0.98
			0.98	1.07
	FEMALE ODDS NORMAL	:DEL	112464:1	64418:1
	FEMALE ODDS NORMAL	:DUP	2084:1	289:1
	MALE ODDS NORMAL	:DUP	9444:1	1920:1
	PROB OF DEVIATION	NORMAL	#####	#####
	FEMALE PROB OF DEVIATION	DELETED	#####	#####
	FEMALE PROB OF DEVIATION	DUP	#####	#####
	MALE PROB OF DEVIATION	DUP	#####	#####

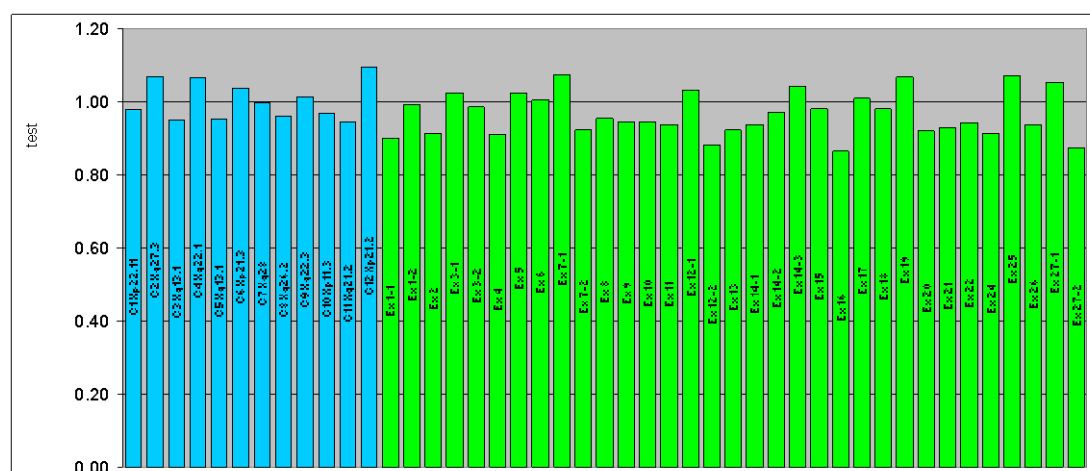
The peak height data generated from a GeneMapper output table is pasted directly into the cells of the raw data sheet. The spreadsheets must have five normal controls in order to function properly and can accept up to 10 test samples. Once the test samples have been added to the raw data sheet, the results may be visualised on the results spreadsheet. This is estimated by measuring the standard deviation of all the test ligation products against each other. The results are presented in a tabular format and the cells are conditionally formatted to highlight deleted/duplicated and aberrant results. The actual settings that have been given for the conditional formats are given in the spreadsheet, for typical ranges see Table 2.5.

Table 2.5: Typical ranges for Dosage Quotients

copy number STATUS	Dosage Quotient
Normal	0.85 – 1.15
Deleted	0.35 – 0.65
Duplicated	1.35 – 1.65
Equivocal	0.65 – 0.85

The spreadsheets also incorporate a graphical representation of the results in the form of a histogram (see Figure 2.5). The probes for each exon of the gene are represented by the green bars in the histogram. The internal reference probes also included in the probemix are represented by the blue bars in the histogram. The *F8* exon numbering used by MRC-Holland adopts the NCBI Refseq exon numbering for *F8*. There is no probe for exon 23 therefore exons 24, 25, 26 and 27 in the MLPA kit relate to exons 23, 24, 25 and 26 of *F8* respectively.

Figure 2.5: Example of a histogram from the P178-B1 *F8* MLPA spreadsheet summarising the mean DQ data. The green bars represent *F8* exons while the blue bars represent internal reference probes. Ratios are calculated by intra-sample normalisation of peak areas to reference probes, and inter-sample normalisation of probe ratios to the corresponding ratios in the reference samples. This particular sample showed normal odds ratios for all ligation products.



The output from the above process could then be interpreted as representing either failed data which required repeat MLPA analysis, or good quality data which could be analysed for evidence of the presence of deletions or duplications associated with the gene undergoing analysis.

CHAPTER 3: RESULTS

At the MDC, a genetic diagnostic service for haemophilia A and VWD has been set-up in line with guidelines from the UKHCDO and is offered for Greater Manchester and the North West. Recent publications have demonstrated the advantages of including MLPA analysis in the genetic diagnosis of haemophilia A and VWD to help identify those with large deletions/duplications and aid carrier diagnosis. MLPA is a technique which was first described in Nucleic Acid Research by Schouten, McElgunn et al. (2002). I have investigated patients with haemophilia A and VWD using this technique in order to identify any mutations which may have gone undetected using conventional analytical methods or confirm deletions detected by an independent method and to introduce it as part of the routine diagnostic service.

3.1 CHOICE OF METHOD

The MLPA kits used in this study were obtained from MRC-Holland who first established MLPA technology. They are the primary producers of MLPA kits and currently offer probe kits for a large number of genes. In this study MLPA probe kits for *F8* and *VWF* were chosen to investigate haemophilia A and VWD respectively.

3.2 MLPA EXPERIMENT OPTIMISATION

As MLPA was a new method to be investigated, it was necessary to carry out a number of optimisation steps before patient samples could be tested. Different parameters such as extraction methods, handling and storage of DNA and experimental set-up have an important role in the quality of MLPA results produced. As part of my investigation into setting up this method, I visited the haematology laboratory at Edinburgh Royal Infirmary who already had this method established as part of their genetic diagnostic service for Haemophilia and VWD.

3.2.1 Results of the Evaporation test

As recommended by MRC-Holland, the use of a thermal cycler with a heated lid (99-105 °C) was essential in order to minimise evaporation losses. A test for

excessive evaporation was performed on the thermocycler to be used for the MLPA experiments. Eight μL of water was incubated overnight at 60 °C in all 8 tubes of two different brands of 0.2 ml 8-strip PCR dome capped tubes (see Table 3.1). The following day, the volume of the water in each tube was compared to see which had the least amount of evaporation. Each of the 8 tubes of the StarLab brand had only $\sim 1 \mu\text{L}$ of water evaporated compared to $\sim 3.5 \mu\text{L}$ in each of the 8 Alpha Labs tubes, therefore it was decided that the StarLab brand of tubes would be used in all experiments.

Table 3.1: Results of the evaporation test comparing two brands of tubes.

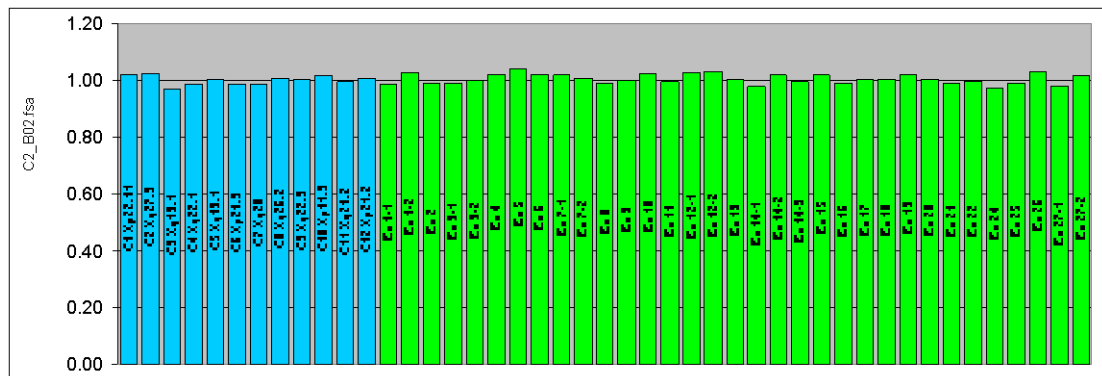
Tube brand	Volume before O/N incubation	Vol. after O/N incubation	Vol. of evaporation
Starlab (I1402-2900)	8 μL	7 μL	1 μL
Alpha Labs (LW2510)	8 μL	4.5 μL	3.5 μL

3.2.2 Evaluation of the MRC-Holland MLPA method using control samples

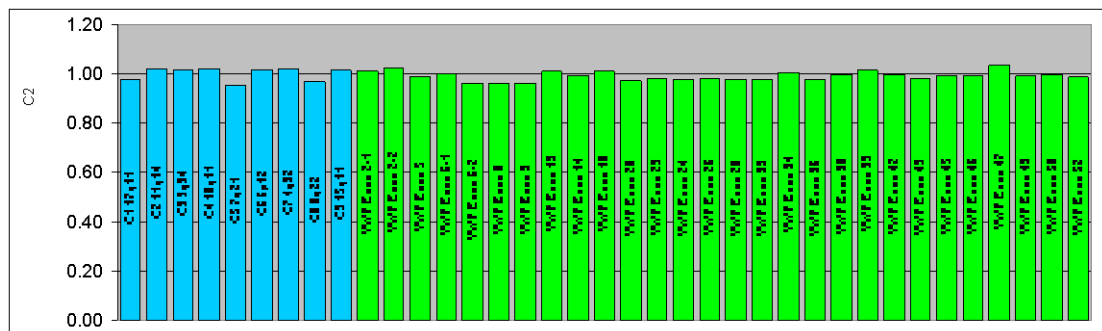
The MRC-Holland MLPA kits for *F8* and *VWF* were first tested using a fresh control sample obtained from a healthy patient. DNA had previously been extracted from the control sample using standard protocols (see Materials and Methods section 2.3). The amount of DNA was measured and diluted to a final concentration of 40 ng/ μL as per the protocol. The control samples were tested according to the standard MLPA procedure and thermal cycler parameters outlined by MRC-Holland (see Materials and Methods section 2.6). Initial results for both kits are shown in Figure 3.1. The histograms showed equal peak heights for all the exons in both kits with a standard deviation of less than 0.1 which demonstrated good results. From this preliminary experiment it was clear that the MLPA technique was a robust method and therefore patient samples could now be tested. It was advised that at least 5 control samples should be included with each MLPA run, therefore it was decided that 7 control samples should be used to cover for any controls that might fail.

Figure 3.1: Initial test control sample histograms for MLPA kits. The height of the columns represents the normalised data obtained in genomic DNA. In a) *F8* and b) *VWF* kit 1, the blue bars represent the internal reference probes and the green bars indicate the probes of each exon. In c) *VWF* kit 2 the first 10 blue bars represent the internal reference probes and the rest represent the probes for each exon. In all the normal controls, the probes are at a height of around 1.0.

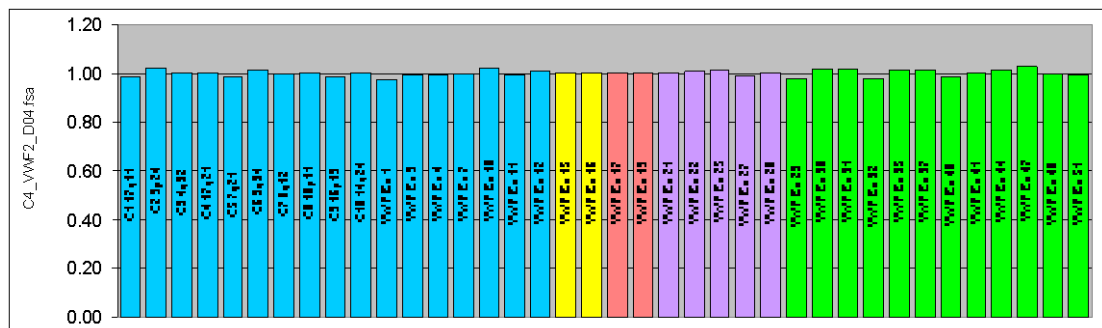
a)



b)



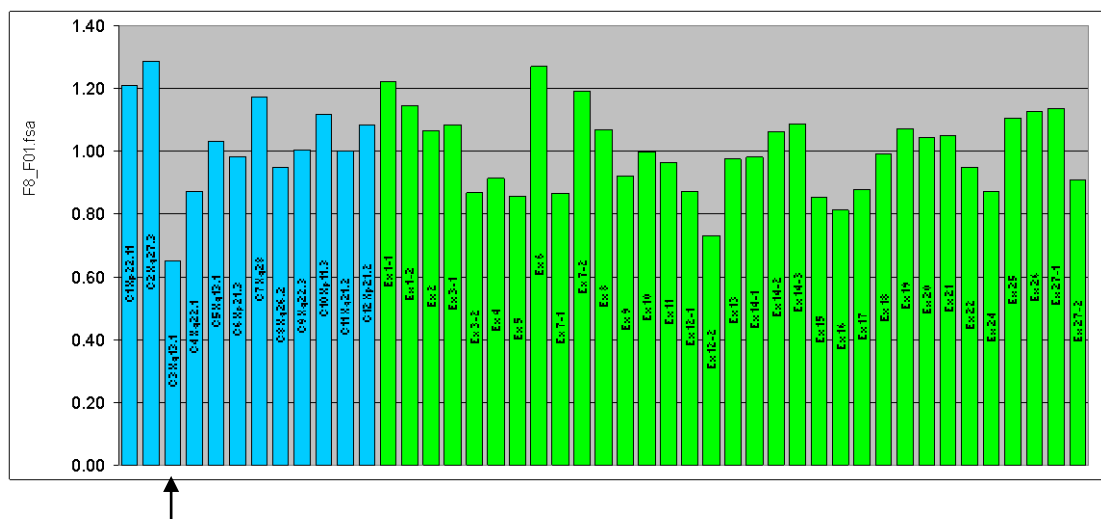
c)



3.2.3 Assessment of sample quality

The extraction methods, handling and storage of DNA have an important role in the quality of MLPA results produced. The DNA used in this project was not from fresh samples and had been stored at -80 °C for a number of years. For best results it is recommended that DNA should be extracted from fresh samples and that a concentration of between 50-250 ng of DNA in a 5 µL final volume is recommended for each MLPA reaction. As the samples had been frozen for a long time, DNA degradation had occurred and so the concentration was often low which would contribute to consistently poor results found for some patients. An example of this can be seen in the histogram below (figure 3.2) for a patient tested with the *F8* MLPA kit. The reference probes in blue are not of a consistent height with the third reference probe, marked by an arrow, considerably lower than the others. This leads to failure of internal normalisation and a standard deviation of more than 0.1 indicating a sample of poor quality where the results cannot be reliably reported.

Figure 3.2: Example of a poor quality sample used for MLPA analysis of *F8* with a standard deviation of >0.1. The third blue reference probe indicated by the arrow is significantly lower than the rest of the reference probes, which leads to failure of intra-normalisation.



3.2.4 Optimisation of the MLPA thermal cycler programme

On examination of the optimised MLPA protocol in use at Edinburgh Royal Infirmary it was apparent that there were slight differences in the conditions of the hybridisation, ligation and PCR reactions compared to the standard MLPA

protocol supplied by MRC-Holland (see Table 3.2 below and Table 2.2 in Materials and methods). These differences may have allowed more consistent MLPA results, especially for samples which repeatedly failed quality control, so they were noted down with the intention of testing these conditions on the thermocycler in Manchester to see if they had an effect.

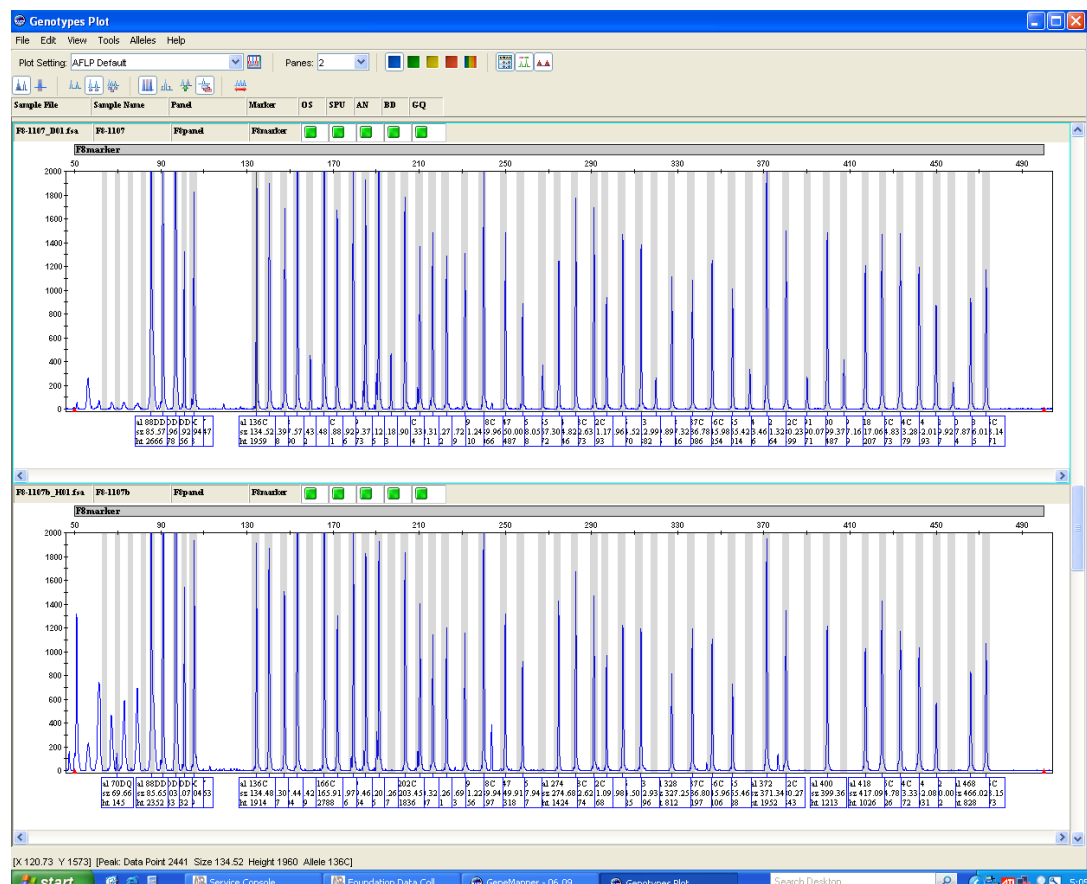
Table 3.2: Thermocycler Program for the MLPA Denaturation, Hybridisation, Ligation and PCR Reactions as per Edinburgh laboratory.

Step	Temperature	Time	No. of cycles
DNA Denaturation	98 °C 22 °C	5 min Pause	Held for 15 min
Hybridisation reaction	95 °C 60 °C	1 min pause	Held for 16 h
Ligation reaction	54 °C 54 °C 98 °C 22 °C	pause 15 min 5 min pause	
PCR reaction	60 °C 95 °C 60 °C 72 °C 72 °C 08 °C	pause 30 s 30 s 60 s 20 min pause	} 33 cycles

To establish whether better quality MLPA results could be obtained using the conditions used in Edinburgh, two experiments were tested in parallel using the two different cycler programmes but with the same set of 6 samples for the *F8* MLPA kit. The PCR products for both sets of experiments were tested together on the fragment analyser to ensure identical analysis conditions. The results from this experiment showed a definite difference between the samples tested with different cycler parameters. Five of the six samples tested showed better results using the standard MRC-Holland protocol compared to the optimised Edinburgh protocol. This was demonstrated by better peak amplification and morphology. An example of this can be seen in Figure 3.3 which shows the genotype plots generated by the GeneMapper® software for one of the samples using the two different parameters. The top plot shows the sample

tested with MRC-Holland parameters and it can be seen that all the peaks to evaluate internal quality control showed good amplification together with peaks for each exon. In comparison the genotype plot below for the same sample tested using Edinburgh parameters showed some non-specific peaks at the beginning of the trace and absence of peaks where the probes had not bound which could potentially indicate the false presence of a deletion. One of the six samples tested using both parameters showed identical results. Following this, when the results were analysed using the NGRL spreadsheet, five out of six of the samples using the Edinburgh parameters showed a standard deviation of more than 0.1 and therefore did not pass quality control indicating that the results were not reliable. From this experiment it was decided that better results were obtained using the standard MRC-Holland protocol and that this method would be used routinely.

Figure 3.3: Genotype plots for a sample tested in parallel to compare the MRC-Holland and Edinburgh parameters. The top plot shows the sample tested with MRC-Holland parameters with good peak amplification of quality control fragments and exons. The bottom plot shows the sample tested with the Edinburgh parameters illustrating non-specific peak amplification at the beginning of the trace and complete absence of peaks for certain exons.

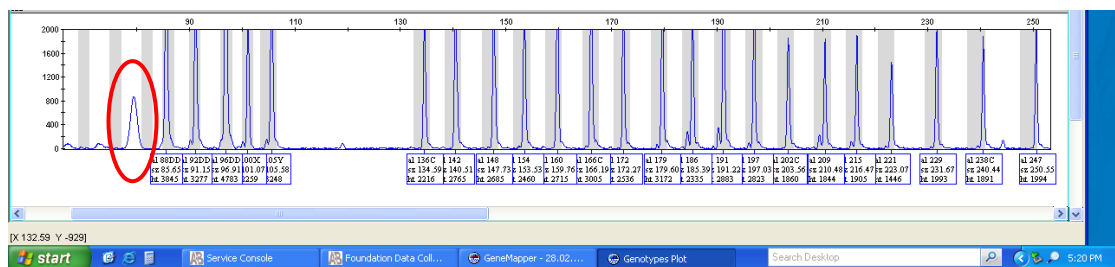


3.2.5 Internal quality controls and Non-specific peaks

As outlined in section 2.7.1 of the Materials and methods, SALSA MLPA probe mixes contain a number of control fragments which identify common problems in the MLPA reaction. The four Q-fragments at positions 64, 70, 76, and 82 nt should be low or invisible when there is sufficient sample DNA present and the successful ligation has occurred as they are present in low quantities and so are outcompeted by the amplicons of the MLPA probes. If the peaks of all four Q-fragments are higher than 1/3 of the height of the 92 nt ligation-dependant control fragment and the MLPA probes this can indicate that the sample DNA was insufficient or failure of the ligation reaction. However if there are only one

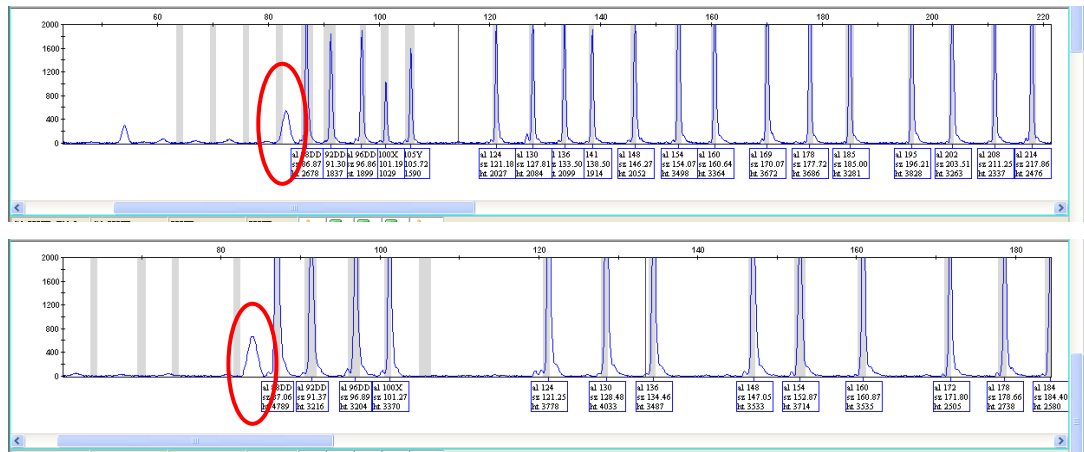
or two increased peaks visible in the 64-82 nt range, this is classed as a non-specific peak rather than a Q-fragment. The presence of non-specific peaks was commonly observed when reviewing the genotype plots for both controls and patient samples. Figure 3.4 shows an example of a non-specific peak (circled in red) observed in the 64-82 nt range for a patient sample tested with the *F8* MLPA kit. This peak would not affect the reliability of the results produced.

Figure 3.4: Example of a non-specific peak in the 64-82 nt range. The peak observed in the 64-82 nt range (circled in red) for a patient sample tested with the *F8* MLPA kit is not considered to be a Q-fragment and therefore would not affect the reliability of the result.



The MLPA probe mixes for *F8* and *VWD* also contain two DNA denaturation fragments (D-fragments) at 88 and 96 nt which show low amplification compared to the 92 nt control fragment if there is incomplete DNA denaturation. A non-specific fragment of ~87 nt was often observed in samples which was close to the 88 nt D-fragment (see Figure 3.5). According to the MRC-Holland MLPA protocol this non-specific fragment at ~87 nt is usually present when using the ABI POP7 polymer for capillary electrophoresis. It can be ignored as it represents non-specific amplification of a long primer-dimer that has the same mobility as one of the quality fragments. If the non-specific peak affected the analysis of the 88 nt D-fragment, only the 96 nt peak was used.

Figure 3.5: Examples of non-specific peaks at ~87 nt observed in two different patient samples. The non specific peaks at ~87 nt (circled in red) observed in two different patient samples tested with VWD MLPA kits 1 (top) and 2 (bottom) which can occur when using the ABI POP7 polymer for capillary electrophoresis.



Non specific peaks were also commonly observed in the 'no DNA' control reaction of some experiments. These were generally found to not have an impact on the overall MLPA results and could be distinguished from any form of contamination. MLPA PCR is more prone to generation of unspecific long amplification products in the absence of sample than a normal PCR as the reaction contains numerous long MLPA probe oligonucleotides, each of which contain a perfect copy of one of the PCR primers.

3.2.6 Bioinformatic analysis

Following experimental set-up and method optimisation, it was necessary to evaluate the software programs that were available for fragment analysis and gene dosage. Initial peaks generated by fragment analysis were evaluated using the ABI GeneMapper® software which was already available in the Molecular Diagnostics Centre. Gene dosage was determined by MLPA spreadsheets devised by the NGRL in Manchester which were available for free download from their website. The gene dosage spreadsheets were first reviewed in April 2011 and it was noted that the spreadsheets for both *F8* and *VWD* did not contain the latest cross reference data for the MLPA probe mixes which had been updated by MRC-Holland in March 2011. Therefore new

spreadsheets were obtained from Sheffield Regional Genetics Service who had updated them for use in their routine diagnostic service.

In the Edinburgh laboratory, an alternative software programme was used called GeneMarker v1.8 (Softgenetics.com) which allowed both raw data to be viewed for evaluation of probe peaks and gene dosage analysis. It was important to establish whether this software was more user friendly and if it was worth considering for future use instead of the combined ABI GeneMapper software and NGRL MLPA spreadsheets. A free one month download version of GeneMarker v1.8 was obtained and previous results were analysed using the software. The previous results analysed by GeneMapper and MLPA spreadsheets were validated using the GeneMarker software. The software proved just as user friendly as the other methods and although it had the added benefit of being able to generate reports, this was not enough to account for the financial cost of the licence needed for use of the full version. Therefore it was decided that the analysis methods that were already in place were sufficient enough for use in the routine service at the MDC.

3.3. RESULTS OF MLPA ANALYSIS IN THE HAEMOPHILIA A PATIENT GROUP

In total, nineteen patients were investigated for large deletion/duplications in *F8* (see Table 3.3). These included haemophilia A patients where no causative mutation had been identified and those with suspected partial gene deletions detected by absence of amplification of one or more exons and relevant female carrier relatives.

Table 3.3: Haemophilia A patients and query carrier relatives investigated by MLPA analysis following routine *F8* screening.

PATIENT	CASE	SEVERITY	FVIII (IU/DL)	INHIBITORS	RESULT PRIOR TO MLPA ANALYSIS
1	Index	?Mild	30	No	None found
2	Index	Mild	11-16	No	None found
3	Index	Mild	13	No	None found
4	Index	Mild	13	No	None found
5	Index	Severe	unknown	Yes	Hemizygous for deletion removing exons 16-22
6	Carrier (Sister of Patient 5)	Severe	Not done	N/A	None found, consistent with non-carriership
7	Index	Severe	1.1	Yes	?hemizygous deletion involving exon 7
8	?Carrier (Mother of Patient 7)	Severe	Not done	N/A	?carrier diagnosis
9	Index	Severe	<1	Yes	Hemizygous for deletion removing exons 15-22
10	?Carrier (Niece of Patient 9)	Severe	Not done	N/A	None found, suggestive of non-carriership
11	Index	Severe	<1	Yes	Hemizygous for deletion removing exons 24-25
12	Index	Severe	unknown	Yes	Hemizygous for deletion removing exon 24
13	Carrier	Severe	72	N/A	Heterozygous for deletion of exon 5
14	Index	Severe	unknown	Yes	Hemizygous for deletion removing exons 7-10
15	Index	Severe	unknown	Yes	Hemizygous for deletion removing exons 7-9
16	Index (Aunt of Patient 17)	Severe	4	No	None found
17	Index	Severe	unknown	unknown	None found
18	Index	Severe	<1	Yes	Hemizygous for c.5644_5657dup14
19	Carrier (Sister of Patient 18)	Severe	40	N/A	Heterozygous for the c.5644_5657dup14

3.3.1 Mild haemophilia A patients with no causative mutation identified

Four patients had been diagnosed with mild haemophilia A. Routine sequencing of *F8* had been previously performed and had failed to identify a candidate mutation, although it has been reported by (Bogdanova, Markoff et al. 2007) that 12% or more of cases classified as mild haemophilia A have no detectable mutation within *F8*.

Patient 1 had a baseline FVIII:C level of 30 IU/dL. As there was no family history of bleeding disorders, it was considered to be a case of sporadic haemophilia. The possibility of type 2N VWD should have been investigated for by VWF-FVIII binding studies or type 2N VWD mutation screening, however no results were available. It was unclear whether FV:C levels had been measured in order to exclude the rare possibility of combined FV/FVIII deficiency.

Patient 2 had reduced FVIII:C levels of between 11-16 IU/dL. Investigations for possible type 2N VWD showed borderline low/normal results for VWF:Ag and VWF:RCo and normal VWF multimers. No FV:C level had been measured to exclude a combined FV/FVIII deficiency. Sequencing of exons 18-20 and 23-27 of *VWF* had also been carried out and had failed to identify any mutations associated with type 2N VWD.

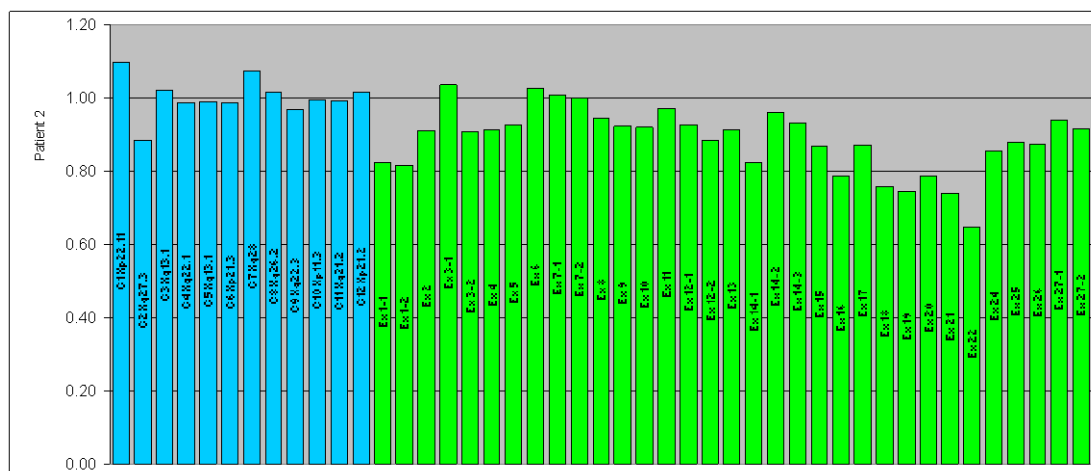
Patients 3 and 4 were brothers who both had baseline FVIII:C levels of 13 IU/dL. Their nephew had also been diagnosed with mild haemophilia A, however no candidate mutations in *F8* had been found in him. Sequencing of exons 18-20 of *VWF* in Patient 3 did not identify any candidate type 2N VWD mutations. Patient 4 had FV:C levels measured at 193 IU/dL, excluding the possibility of a FV/FVIII deficiency.

MLPA analysis carried out on this group of haemophilia patients gave normal results and did not identify any large deletions or duplications within *F8*. The analysis was performed at least twice on different occasions for each patient, thereby confirming the initial result. The histograms show that the probes for each exon were amplified (see Figure 3.6), indicating that no large deletions were present. The heights of the internal reference probes (blue bars) were

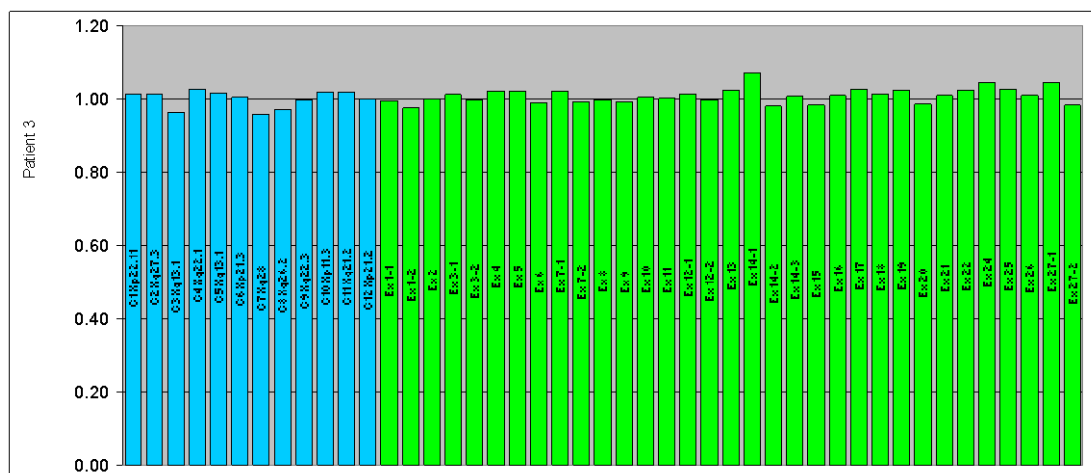
Figure 3.6: MLPA Histograms for Haemophilia A Patients 1, 2, 3 and 4. The histograms for **a)** Patient 1, **b)** Patient 2, **c)** Patient 3 and **d)** Patient 4 illustrated normal dosage quotients for all exons of *F8* (green bars). The histogram for Patient 1 illustrates why the sample gave a standard deviation of >0.1 due to the low peak height of one of the internal reference probes (blue bar indicated by an arrow).

Gene Construct	Patent Count
C18a22.1	1.12
C18a22.2	1.32
C18a22.3	0.72
C33aB.1	0.94
C48a21.1	0.92
C54a21.1	1.02
C63a21.2	1.07
C79a22.1	0.88
C83a22.2	0.98
C84a21.3	1.01
C89a21.2	1.07
C118a21.2	1.14
C124a21.2	1.12
Ex1-1	1.12
Ex1-2	1.13
Ex2	1.05
Ex3-1	0.83
Ex3-2	1.03
Ex4	0.95
Ex5	1.01
Ex6	1.18
Ex7-1	1.05
Ex7-2	1.20
Ex8	1.23
Ex9	1.08
Ex10	1.11
Ex11	0.94
Ex12-1	0.98
Ex12-2	0.98
Ex13	0.93
Ex14-1	0.94
Ex14-2	1.19
Ex14-3	1.20
Ex15	1.09
Ex16	0.92
Ex17	0.93
Ex18	1.14
Ex19	0.94
Ex20	1.16
Ex21	1.13
Ex22	0.88
Ex24	0.83
Ex25	0.87
Ex26	0.98
Ex27-1	0.87
Ex27-2	0.96

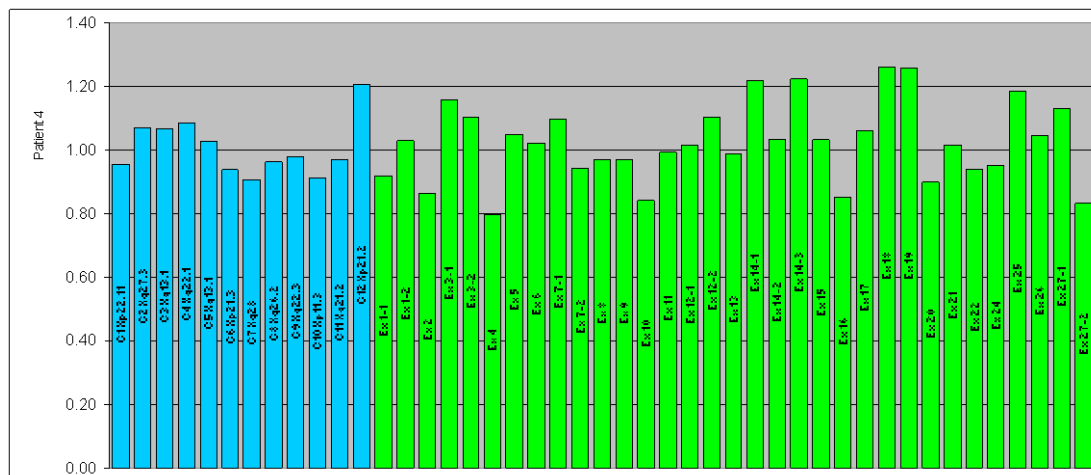
b)



c)



d)



3.3.2 Severe haemophilia A patients with large deletions and carrier relatives

Nine families were investigated for large deletion mutations. Those with suspected gene deletions had been discovered through routine screening by failure of one or more exons to amplify by PCR. Further mutation analysis had been performed for some of the families using primer pairs picked from the standard stock to amplify either side of the expected deletion breakpoint. These primer pairs were utilised in the ABgene Long PCR protocol to amplify across the breakpoint, with deletions being visible on gel electrophoresis as smaller than expected fragments. The TaKaRa protocol for Long PCR was also utilised to amplify fragments which were too large for the ABgene protocol (see Introduction section 1.19.1.4). MLPA analysis was performed on all nine families to confirm the presence or absence of a deletion or duplication. All MLPA results obtained from these families were of good quality (standard deviation of <0.1) and therefore could be reported reliably (see Materials and methods section 2.8).

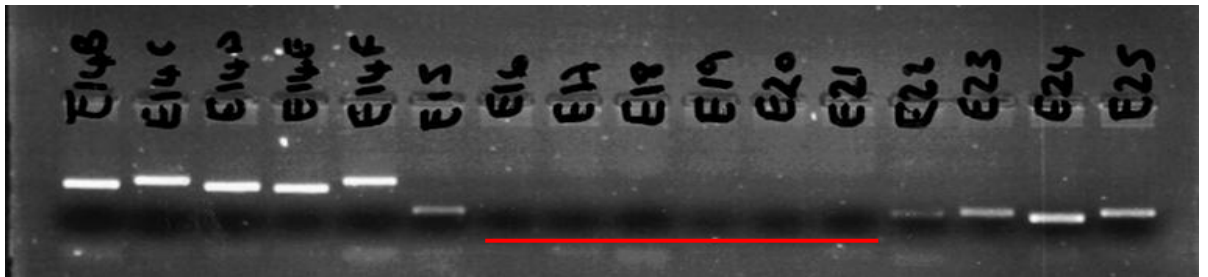
3.3.2.1 Patients 5 and 6

Patient 5 is a severe haemophilia A affected male with an unreported FVIII:C level and a high titre FVIII inhibitor. He had also been reported to be normal for the intron 1 and intron 22 inversions. Mutation analysis had been previously carried out with the aim of identifying the underlying candidate mutation associated with haemophilia A in this family. Amplification of the essential regions of the gene from the genomic DNA had failed to amplify exons 16 to 21 (Figure 3.7a). Subsequent long PCR utilising a primer pair designed to amplify exons 15 through to 22 produced a ~3.25 kb fragment instead of the expected ~11 kb fragment indicating the presence of a partial deletion of *F8* between introns 15-21. Limited DNA sequencing of the ~3.25 kb fragment across the breakpoint had confirmed the identity of this product and indicated that the deletion was likely to be a 7356bp fragment removing exons 16 to 21 (Figure 3.7b). This deletion was highly likely to account for the severe haemophilia A phenotype seen in the patient. Long PCR encompassing exons 15 and 22 could be applied to carrier diagnosis of female members of this family.

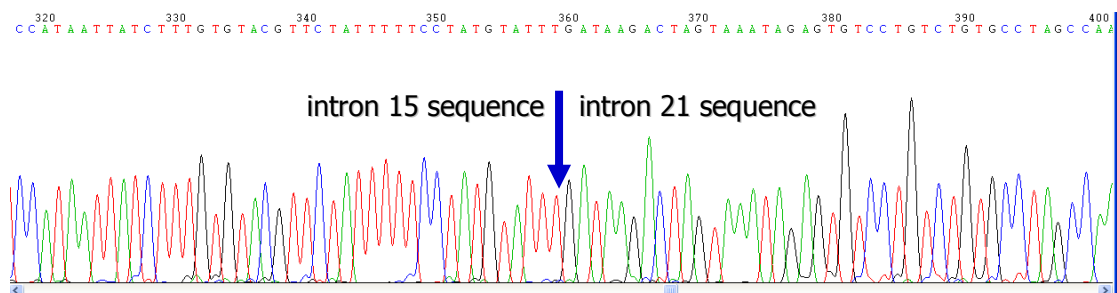
Figure 3.7: Results from previous mutation analysis of Patient 5.

a) Image of the agarose gel where exons 16-21 failed to amplify following standard PCR amplification (underlined in red) **b)** Sequence trace for the partial gene deletion (g.116547_123812del7356) showing the exon 15 breakpoint (indicated by the blue arrow).

a)



b)

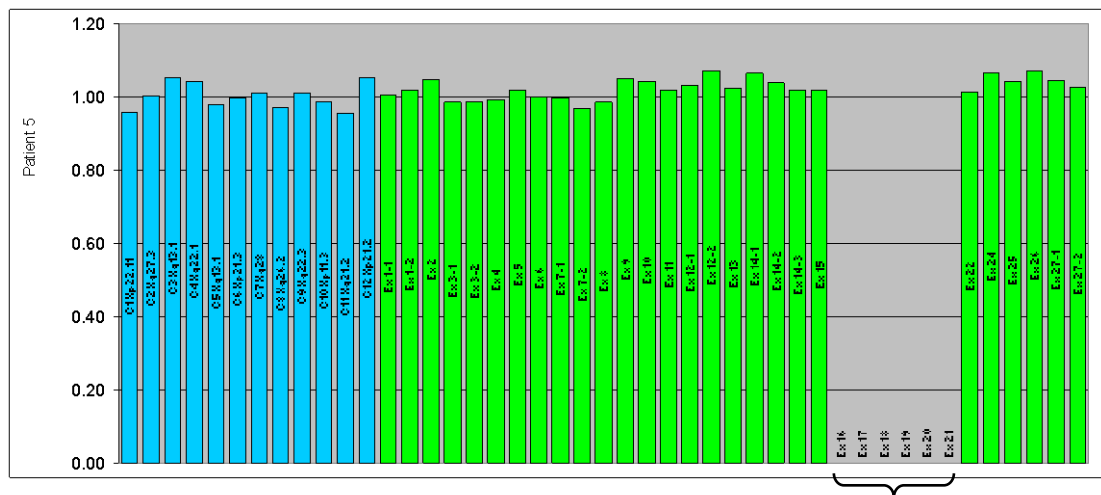


Carrier diagnosis had also been carried out on the sister (Patient 6) of Patient 5 to investigate for the presence of this deletion by long PCR. Patient 6 was found to have only amplified the normal fragment (~11 kb compared to the ~3.25 kb fragment representing the deletion) and partial DNA sequencing confirmed the identity of the ~11 kb normal fragment. This genotype was consistent with a diagnosis of non-carriership.

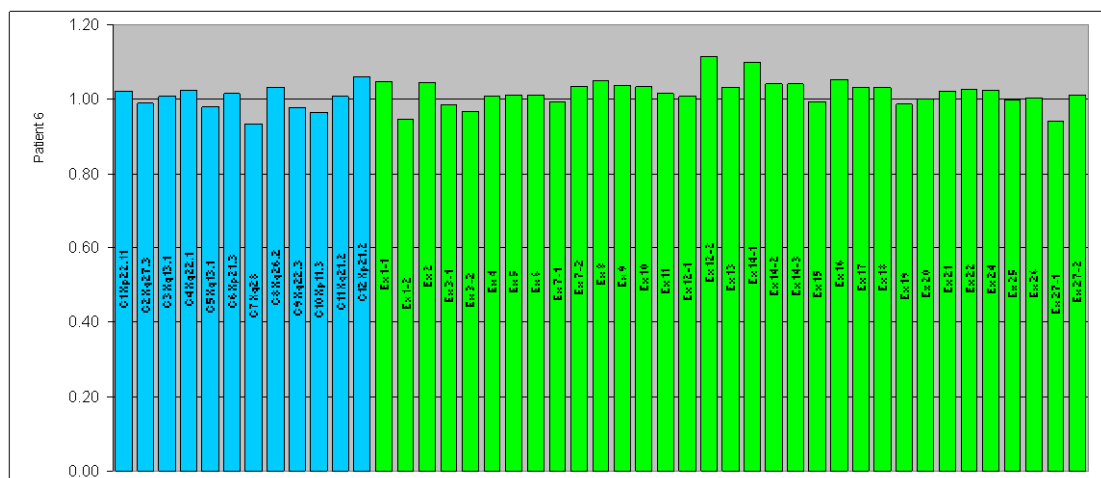
MLPA analysis for both patients confirmed the results that had been obtained using long PCR and sequencing. For Patient 5 the large deletion of exons 16-21 can be clearly seen in the MLPA histogram (Figure 3.8a) where the peaks are absent for exons 16-21. The genotype is consistent with a diagnosis of severe haemophilia A. Patient 6 (Figure 3.8b) shows normal dosage quotients for all exons in that sample confirming non-carriership of the partial gene deletion.

Figure 3.8: MLPA Histograms for Patients 5 and 6. The histogram for **a)** Patient 5 illustrates the hemizygous deletion of exons 16-21 (indicated by the bracket) where the green bars representing these exons are absent. The histogram for **b)** Patient 6 shows normal dosage quotient data for all the exons of *F8* indicating that the deletion is not present.

a)



b)



3.3.2.2 Patients 7 and 8

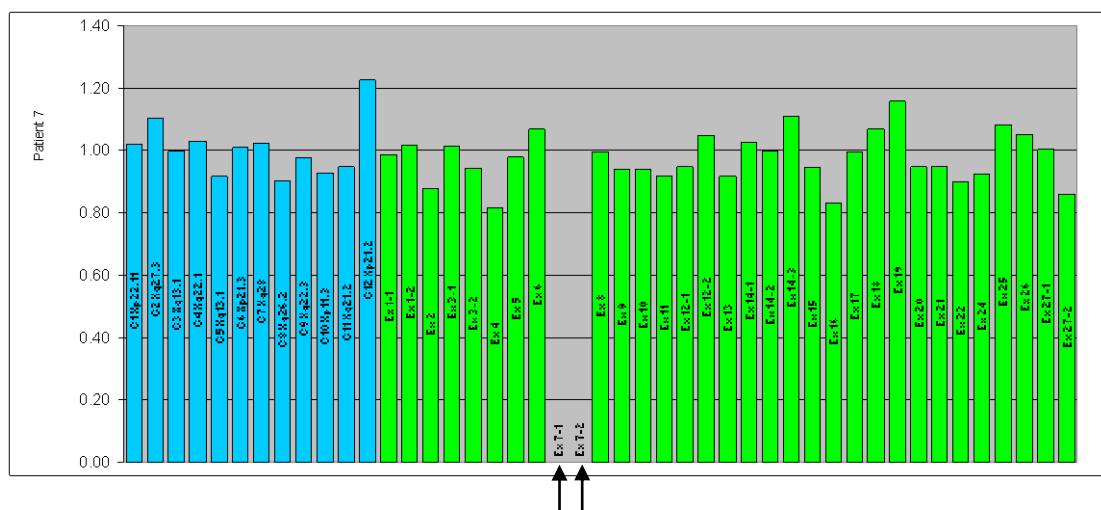
This family was initially investigated as the mother (Patient 8) of the index case (Patient 7) was pregnant. Patient 7 is a severe haemophiliac with a baseline FVIII:C level of 1.1 IU/dL and a FVIII inhibitor. Routine screening of *F8* by sequencing had been normal apart from the inability to amplify exon 7. Repeat analysis had still failed to amplify exon 7 which provided indirect evidence that a partial gene deletion involving this region may be present. As attempts to

map this region failed a possible genotype of a suspected hemizygous deletion involving exon 7 was reported. As the precise nature of the mutation had not been mapped it was not possible to provide a female carrier diagnosis for the mother as the deletion would have been masked by the normal allele on the other chromosome. The mother later gave birth to a female.

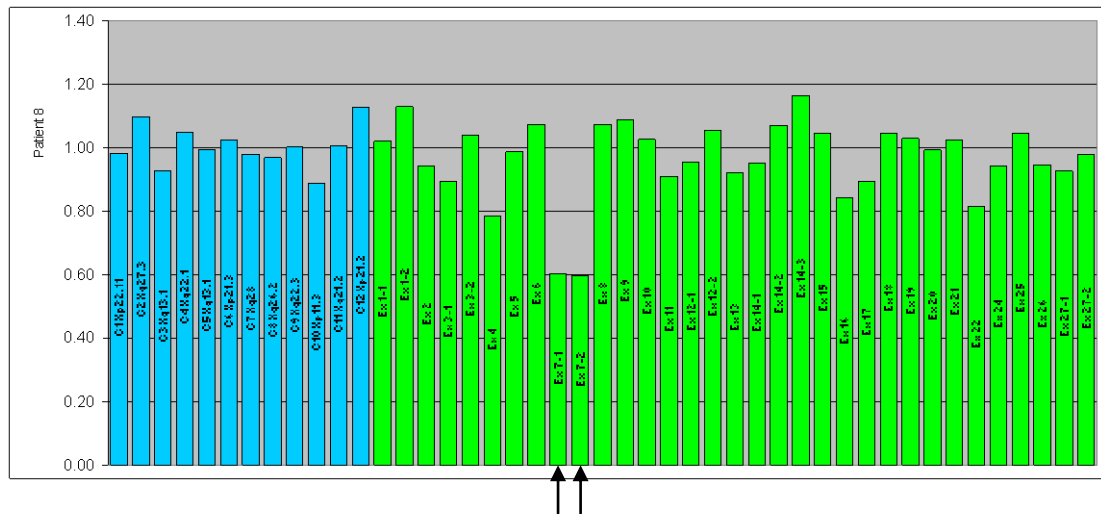
MLPA analysis of both of these patients confirmed the suspected deletion involving exon 7 as the underlying genetic cause of the haemophilia. The histogram for Patient 7 (Figure 3.9a) shows that both peaks that should be amplified for exon 7 are completely missing indicating that this exon has been deleted. The histogram for Patient 8 (Figure 3.9b) shows that both peaks for exon 7 are half the height of the control peaks indicating that this is a heterozygous deletion involving exon 7, which confirms the carrier status of his mother.

Figure 3.9: MLPA Histograms for Patients 7 and 8. The histogram for **a)** Patient 7 illustrates the hemizygous deletion of exon 7 (indicated by the arrows) where the probes representing exon 7 are absent. The histogram for **b)** Patient 8 shows the presence of a heterozygous deletion of exon 7 where the probes representing exon 7 (indicated by arrows) are half the height of the internal reference probes (blue bars).

a)



b)



3.3.2.3 Patients 9 and 10

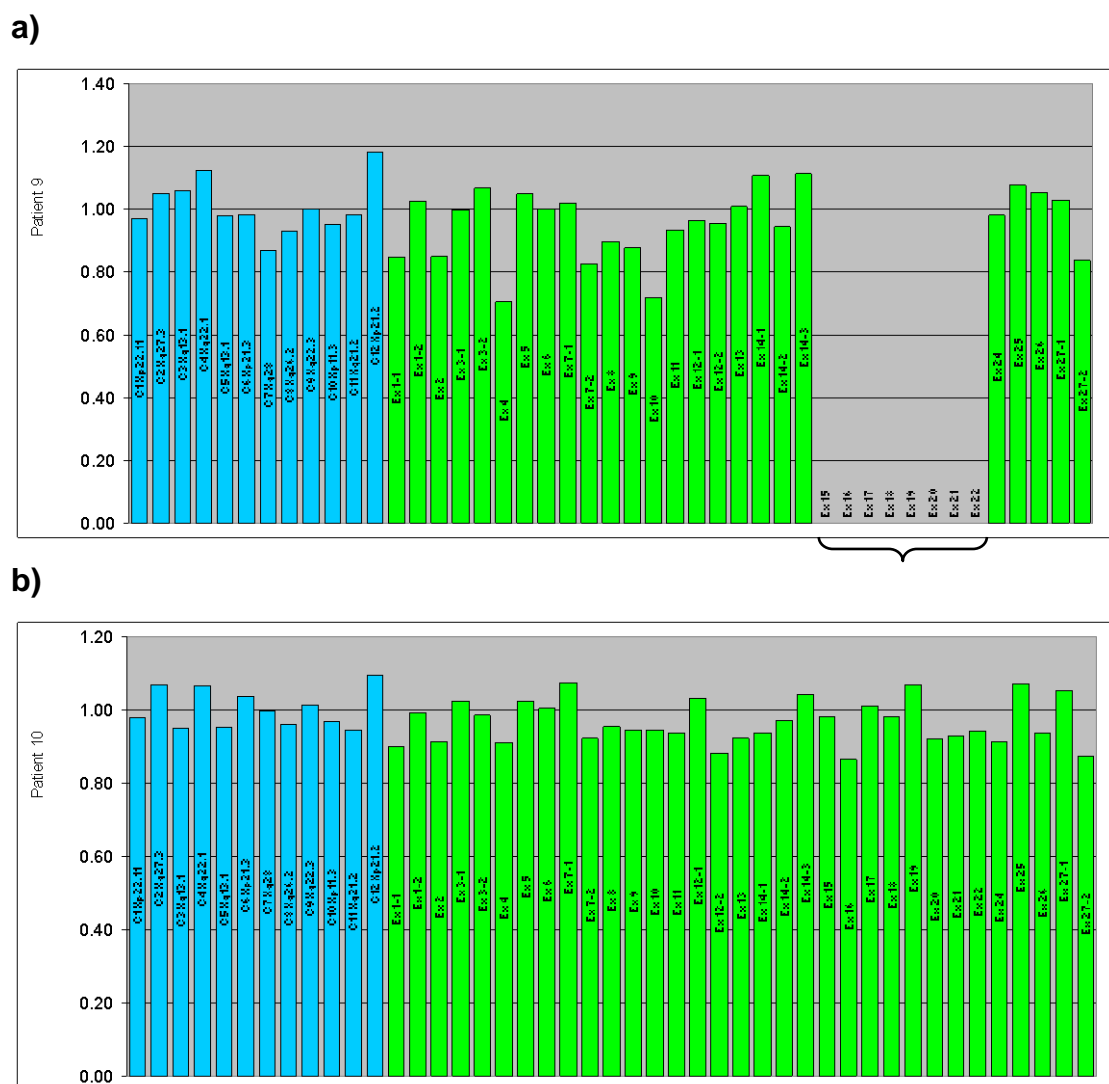
The index case for this family (Patient 9) had severe haemophilia A (FVIII:C level <1 IU/dL) and a FVIII inhibitor. This patient was previously found to be hemizygous for a large deletion spanning exons 15-22 following DNA sequencing. Long PCR had also been performed using primers for exons 14-23 which produced a 10 kb truncated fragment indicating a partial deletion, however the precise deletion boundaries had not been mapped. This exon 15-22 deletion has been reported, as would be expected, to be associated with a severe phenotype. It was suggested that further analysis should be performed in order to map the boundaries of this deletion breakpoint prior to developing a direct detection assay that could be used in female carrier studies within this family.

Carrier diagnosis had also been carried out on the niece (Patient 10) of Patient 9 to identify if she was heterozygous for the partial gene deletion. Long PCR had been performed using the same primer pair but did not detect the truncated product. This negative finding was suggestive of non-carriership, however the result needed to be corroborated by development of a direct screening assay for this test when the deletion was fully mapped.

MLPA analysis of both patients confirmed the results that had been obtained through sequencing and long PCR. The Haemophilia Centre in Sheffield had

previously confirmed these results using MLPA, thereby this was an independent validation of their results. It can be clearly seen in the histogram for Patient 9 (Figure 3.10a) that the peaks associated with exons 15-22 are completely absent indicating a partial gene deletion. The histogram for Patient 10 (Figure 3.10b) shows normal dosage quotients for all exons in that sample confirming non-carriership of the partial gene deletion.

Figure 3.10: MLPA Histograms for Patients 9 and 10. The histogram for **a)** Patient 9 illustrates the hemizygous deletion of exons 15-22 (indicated by the bracket) where the green bars representing these exons are absent. The histogram for **b)** Patient 10 shows normal dosage quotient data for all the exons of *F8* indicating that the deletion is not present.

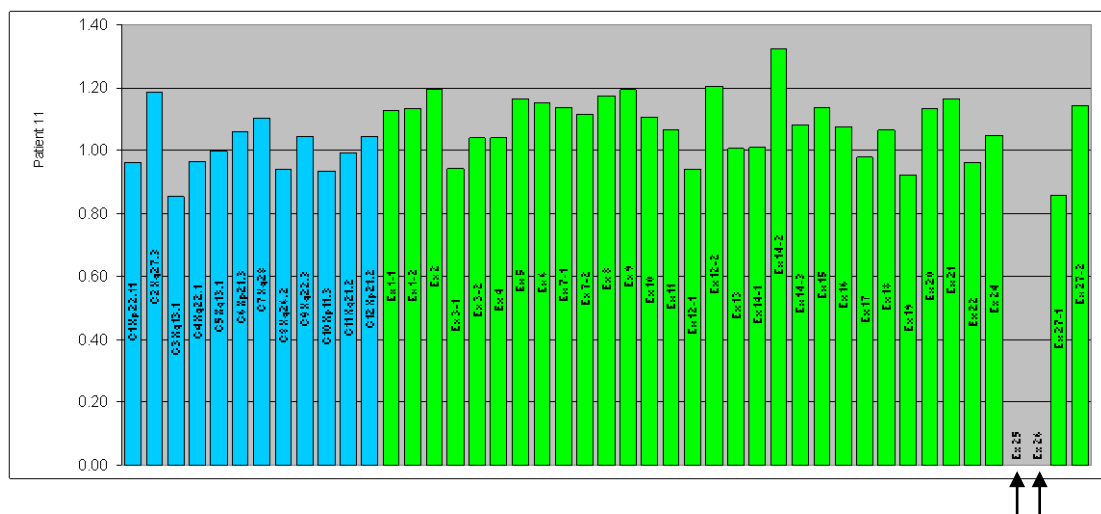


3.3.2.4 Patient 11

Patient 11 was the proband for this family with severe haemophilia A (FVIII:C <1 IU/dL) and a high titre FVIII inhibitor. Previous screening for the intron 1 and 22 inversions had shown normal results and standard PCR amplification of the essential regions of *F8* failed to amplify exons 24-25. Therefore this patient was diagnosed to be hemizygous for a candidate deletion removing exons 24-25 of *F8* which is consistent with severe haemophilia A phenotype. It was suggested that further analysis should be performed in order to map the boundaries of the deletion breakpoint to allow the development of direct detection assay for female carrier studies within this family.

MLPA analysis carried out on the same DNA sample from this patient confirmed the presence of a deletion involving exons 24-25. The MLPA histogram for this patient (Figure 3.11) shows the deletion encompassing exons 24-25, which is consistent with a severe haemophilia A phenotype.

Figure 3.11: MLPA Histogram for Patient 11. The histogram for Patient 11 illustrates the hemizygous deletion of exons 24-25 (indicated by the arrows) where the probes representing these exons (MRC-Holland exon 25-26) are absent.

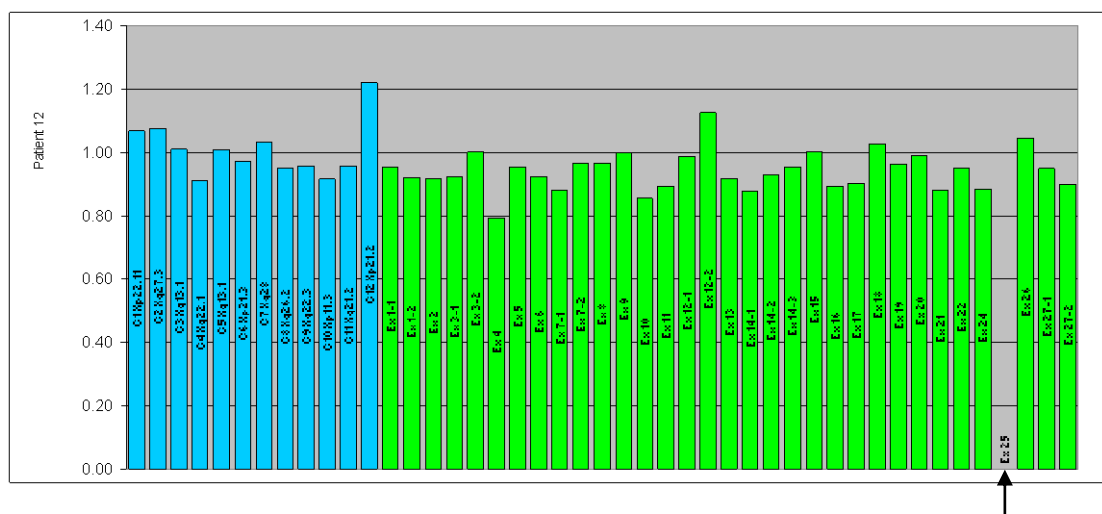


3.3.2.5 Patient 12

Patient 12 was a severe haemophilia A affected male with a FVIII inhibitor. Routine screening for the intron 1 and intron 22 inversions showed normal results. Sequence analysis identified a deletion mutation removing exon 24 of *F8*. Subsequent deletion breakpoint analysis showed that the deletion encompassed 1328 bases of genomic DNA. The 5' breakpoint lies within intron 23 of *F8* and the 3' breakpoint lies within intron 24 (c.6575-262_6724-192del). This deletion predicts a frame shift mutation affecting amino acid valine at position 2242 in the FVIII protein creating a downstream premature stop codon p.(Val2242Serfs*3). There is one entry on the HAMSTERS database describing a similar mutation (Nakaya, Hsu et al. 2004) in association with inhibitor development in severe haemophilia A. Therefore his genotype was reported as hemizygous for the c.6575-262_6724-192del mutation. A multiplex PCR was developed specifically to identify this mutation allowing genetic carrier diagnosis to be offered to female family members if required.

MLPA analysis confirmed the deletion of exon 24 (MRC-Holland exon 25) in this sample, which can be clearly seen in the MLPA histogram (Figure 3.12).

Figure 3.12: MLPA Histogram for Patient 12. The histogram for Patient 12 illustrates the hemizygous deletion of exon 24 (indicated by the arrow) where the probe representing this exon (MRC-Holland exon 25) is absent.



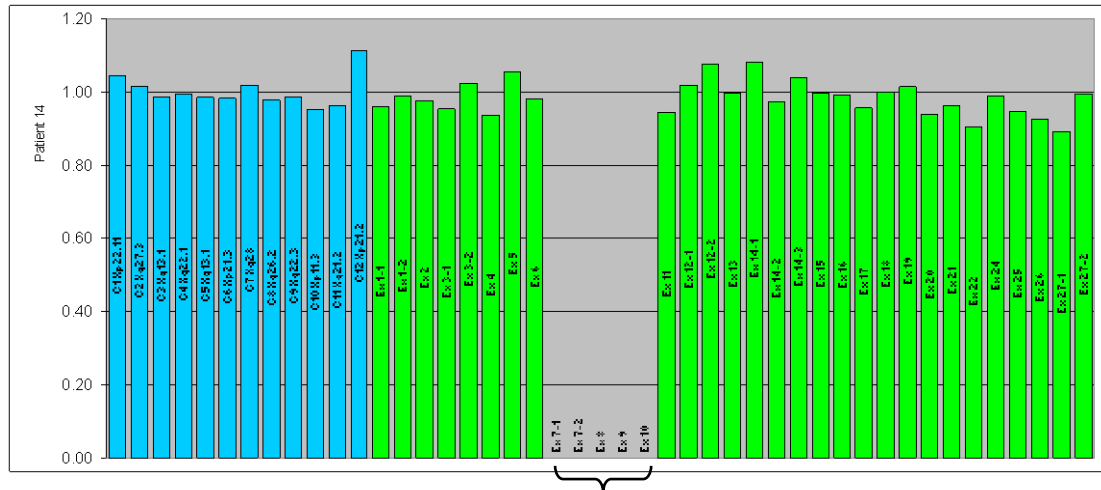
3.3.2.6 Patient 13

Patient 13 was a mother of a severe haemophiliac in whom a partial *F8* deletion mutation has been identified. The index case had a FVIII:C level of <1 IU/dL and was normal for the intron 1 and intron 22 inversions. Mutation analysis had been previously carried out on the index case to identify the underlying candidate mutation associated with haemophilia A in the family. Classic PCR of the essential regions of *F8* from the genomic DNA failed to amplify exon 5. Subsequent long PCR using a primer pair to amplify from exon 4 to 6 produced a truncated fragment of approx 6.5 kb, rather than the expected 8.5 kb product, consistent with a deletion of exon 5. Therefore the index case was found to be hemizygous for the exon 5 deletion which is consistent with a severe haemophilia A phenotype.

Specific analysis for the familial deletion mutation had been previously carried out on the mother who had a FVIII:C level of 72 IU/dL. Long PCR amplification of exons 4-6 (utilising the same primer pair which detected the ~6.5 kb truncated product) showed compound heterozygosity for the ~6.5 kb truncated product and the ~8.5 kb wild type product, which was consistent with carriership for the deletion. The deletion still remains to be fully mapped therefore it was suggested that the result should be further corroborated by the development of a direct screening assay for this test when this has been carried out.

MLPA analysis was only carried out on a DNA sample from the mother as the sample from the index case had been depleted. This confirmed the heterozygous deletion of exon 5 in the mother shown in the MLPA histogram (Figure 3.13) which is consistent with carriership.

Figure 3.14: MLPA Histogram for Patient 14. The histogram for Patient 14 illustrates the hemizygous deletion of exons 7-10 (indicated by the bracket) where the probe representing these exons are absent.

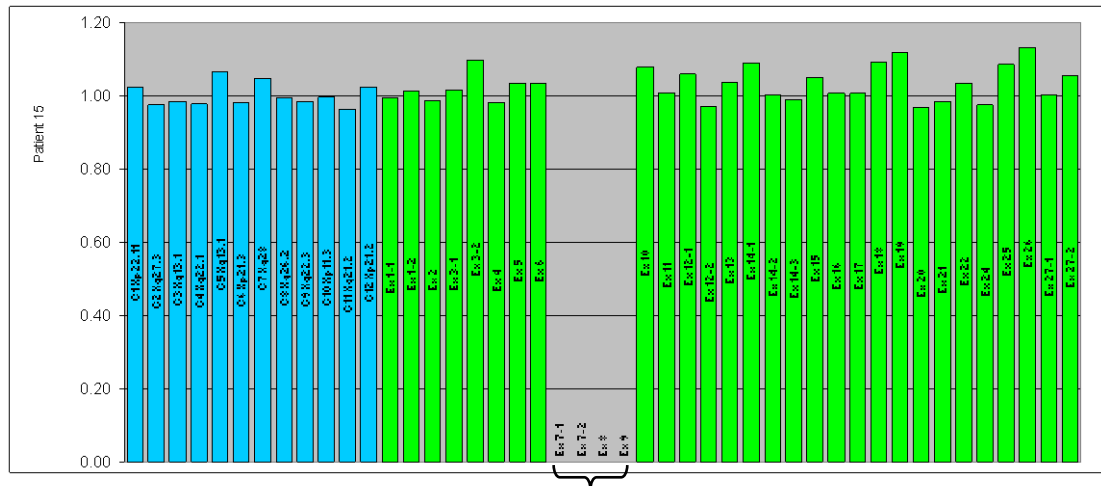


3.3.2.8 Patient 15

Patient 15 was a severe Haemophilia A affected male with a high titre FVIII inhibitor. He had an unreported FVIII:C baseline level and was normal for the intron 22 inversion. He was reported to have a deletion encompassing exons 7 to 9 of *F8*. Previous PCR amplification of the essential regions of *F8* failed to amplify exons 7-9 inclusive which corroborated this earlier observation. A long PCR protocol using a primer pair designed to amplify exons 6 through to 10 produced a 14 kb fragment instead of the expected ~24 kb fragment, indicating that Patient 15 had a deletion of around 10 kb. The deletion removed exons 7 to 9 and was consistent with the haemophilia A phenotype seen in the patient.

MLPA analysis confirmed the deletion of exons 7-9 in this sample, shown in the MLPA histogram by absence of the respective peaks (Figure 3.15).

Figure 3.15: MLPA Histogram for Patient 15. The histogram for Patient 15 illustrates the hemizygous deletion of exons 7-9 (indicated by the bracket) where the probes representing these exons are absent.



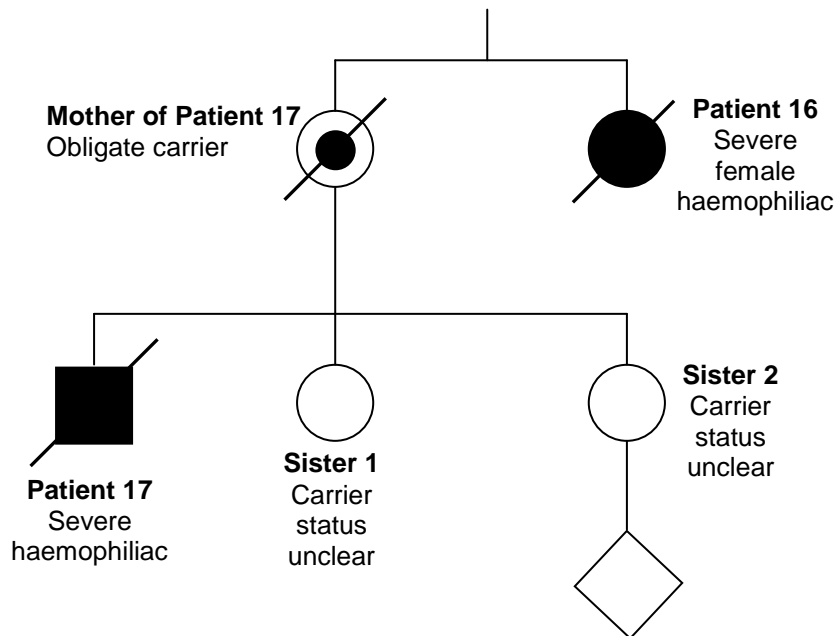
3.3.2.9 Patients 16 and 17

Patient 16 was a deceased female haemophiliac with a FVIII:C level of 4% (with a phenotype similar to a severe haemophiliac) and normal intron 22 inversion results. Her father was a haemophiliac (now deceased) therefore she was an obligate carrier of haemophilia A which was probably inherited from the paternal side of the family. Her mother was also an obligate carrier as she had a brother who probably had severe haemophilia and died of a haemorrhage. The sister of Patient 16 also had a severe haemophiliac son (Patient 17) (see Figure 3.16 for family tree). Therefore it is thought that Patient 16 may have also inherited a haemophilia associated allele from the maternal side of the family (of which there is a 50% chance) which would make her compound heterozygous for two haemophilia associated alleles. However there was no hard evidence to support this supposition, although it may have explained the low FVIII levels seen in Patient 16.

Patient 17 (now deceased) was the nephew of Patient 16 and was an affected male with severe haemophilia A, although there was no baseline level reported. He had also been reported to be normal for the intron 1 and intron 22 inversions. Sequence analysis within the coding regions of *F8* failed to identify any mutations. MLPA analysis carried out by Newcastle Haemophilia Centre

failed to identify any duplications or deletions. Therefore the causative mutation in this patient remained unidentified.

Figure 3.16: Family tree for Patients 16 and 17

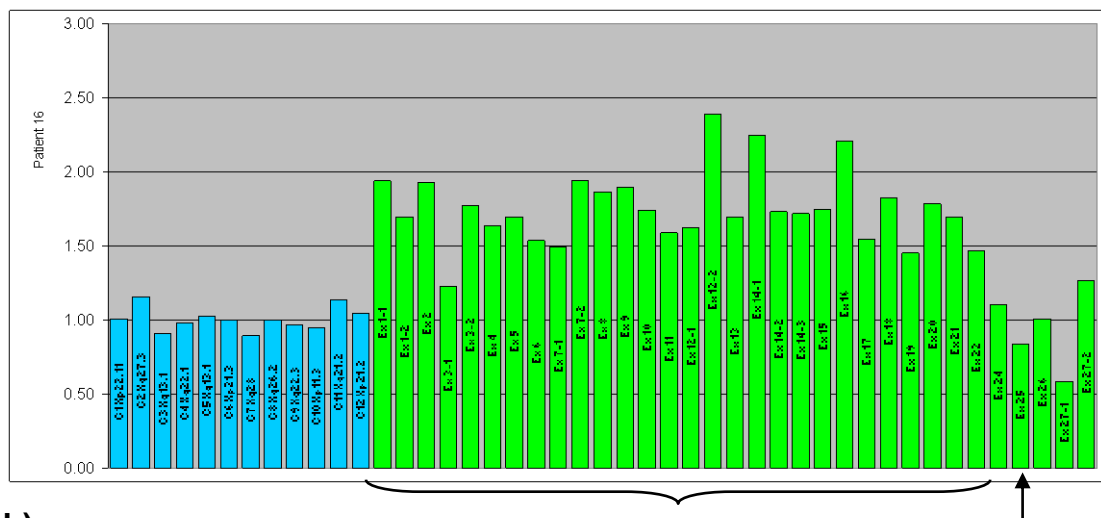


MLPA analysis for Patient 16 suggested a large duplication of exons 1-22 combined with a heterozygous deletion of exon 24. The sample analysed for this patient was very old and the data quality for individual probes was variable as can be seen in Figure 3.17a, however the internal reference probes in blue were fine. The duplication is illustrated in the MLPA histogram (Figure 3.17a) by the peak heights of the probes for exons 1-22 which are twice the height of the internal reference probes. The heterozygous deletion is illustrated by the reduced probe height for exon 24 which is half the height of the internal reference probes.

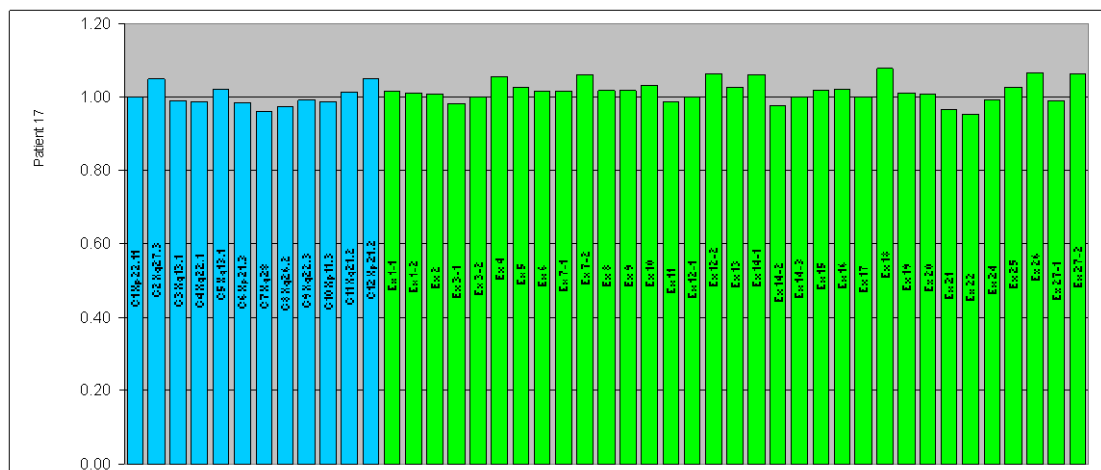
MLPA analysis for Patient 17 (Figure 3.17b) showed normal dosage quotients for all exons within *F8*, therefore the underlying genetic mechanism for the haemophilia within this patient remains unidentified. This also was an independent validation of the results obtained from Newcastle.

Figure 3.17: MLPA Histograms for Patients 16 and 17. The histogram for **a)** Patient 16 illustrates the presence of a possible combined duplication and deletion event. The duplication of exons 1-22 (underlined by the bracket) is indicated by the probes representing these exons which are twice the height of the internal reference probes (blue bars). The heterozygous deletion of exon 24 (indicated by the arrow) is half the height of the internal reference probes. The histogram for **b)** Patient 17 shows normal probe heights for all the exons of *F8* indicating that neither the duplication nor deletion is present.

a)



b)



3.3.2.10 Patients 18 and 19

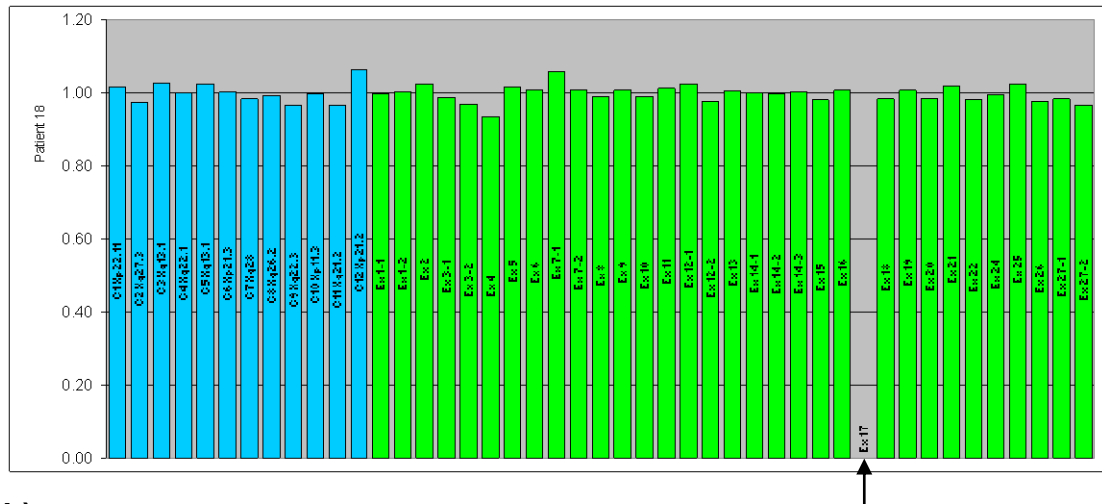
Patient 18 was a severe haemophilia affected male with a FVIII:C level of <1 IU/dL and a FVIII inhibitor. Mutation analysis of *F8* had been previously carried out in this patient and had identified a DNA 14 base duplication c.5644_5657dup14 within exon 17. This nucleotide sequence duplication predicted a frame shift mutation affecting amino acid histidine at position 1886 in the FVIII protein, creating a downstream premature stop codon p.(Gly1887*). The c.5644_5657dup14 nucleotide duplication was not listed on the *F8* mutation databases and its absence had not been confirmed in an unaffected male from the maternal side of the family. However, given the nature of the mutation it was highly likely to account for the haemophilia A phenotype present in the family. Patient 18 was hemizygous for the *F8* c.5644_5657dup14 mutation which was consistent with his diagnosis of severe haemophilia A.

Carrier diagnosis had also been carried out on the sister (Patient 19) of Patient 18 to investigate for the presence of the familial duplication by means of direct DNA sequencing. Patient 19 had normal FVIII:C levels of 40 IU/dL. She was found to be heterozygous for the *F8* c.5644_5657dup14 mutation which was consistent with a diagnosis of carriership for haemophilia A.

MLPA analysis for Patient 18 showed drop out of the probe for exon 17 due to the 14 bp duplication. This is illustrated in the MLPA histogram (Figure 3.18a) by complete absence of the peak for exon 17. MLPA analysis for Patient 19 (Figure 3.18b) showed the heterozygous 14 bp duplication by the reduced peak height of exon 17 which was half the height of the internal reference probes. The presence of a mutation or polymorphism in the sequence detected by a probe can cause a reduction in the relative peak height. Figure 3.19 shows the gene transcript for exon 17 and the location of the *F8* MLPA probe highlighted in pink and green. The location of the c.5644_5657dup14 mutation is highlighted in purple which is present within the probe sequence. Therefore the presence of this mutation will cause a reduction in peak height of the probe.

Figure 3.18: MLPA Histograms for Patients 18 and 19. The histogram for **a)** Patient 18 illustrates the probe drop out for exon 17 (indicated by the arrow) due to the hemizygous 14 bp duplication. The histogram for **b)** Patient 19 shows a reduced probe height for exon 17 caused by the heterozygous 14 bp duplication.

a)



b)

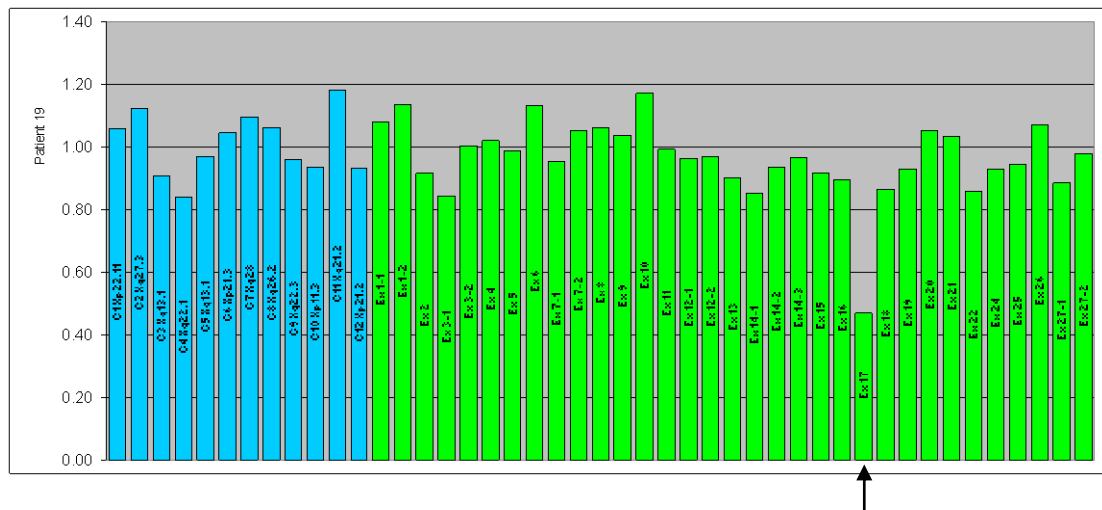


Figure 3.19: F8 transcript for exon 17. The transcript shows the location of the exon 17 probe (highlighted in pink and green) and the location of the c.5644_5657dup14 mutation (highlighted in purple) which can lead to complete absence or reduction in height of the probe.

Exon 17

5701 ^Y ^M CCCACTAAAGATGAGTTTGACTGCAAAG^YCTGGGCTTATTTCTCTGATGTTGACCTTGAA
5530 CCCACTAAAGATGAGTTTGACTGCAAAGCCTGGGCTTATTTCTCTGATGTTGACCTTGAA
1844 =P==T=-K--D--E--F--D--C--K--A=-W--A--Y--F--S--D--V--D--L--E-

Exon 18

5761 ^D ^R AAAGATGTGC^RACTCAGGCCTGATTGGAC^SCCTTCTGGTCTGCCACACTAACACACTGAAC
5590 AAAGATGTGCACTCAGGCCTGATTGGACCCCTTCTGGTCTGCCACACTAACACACTGAAC
1864 -K=-D=-V=-H=-S--G--L--I--G--P=-L--L--V--C--H--T--N--T--L--N-

5821 CCTGCTCATGGGAGACAAGTGACAGTACAGGAATTTGCTCTGTTTTTCACCATCTTTGAT
5650 CTTGCTCA^RGGGAGACAAGTGACAGTACAGGAATTTGCTCTGTTTTTCACCATCTTTGAT
1884 -P--A--H--G--R--Q--V--T--V--Q--E--F--A--L--F--F--T--I--F--D-

5881 ^R GAGACCAAAGCTGGTACTTCACTGAAAATATGGAAAGAACTGCAGGGCTCCCTGCAAT^{SS}
5710 GAGACCAAAGCTGGTACTTCACTGAAAATATGGAAAGAACTGCAGGGCTCCCTGCAAT
1904 =E=-T--K--S--W--Y--F--T--E--N--M--E--R--N--C--R--A--P--C=-N-

Exon 18

5941 ATCCAGATGGAAGATCCCACTTTTAAAGAGAATTATCGCTTCCATG^RCAATCAATGGCTAC
5770 ATCCAGATGGAAGATCCCACTTTTAAAGAGAATTATCGCTTCCATGCAATCAATGGCTAC
1924 -I--Q--M--E--D--P--T--F--K--E--N--Y--R--F--H--A--I--N=-G--Y-

3.4 RESULTS OF MLPA ANALYSIS IN THE TYPE 3 VWD PATIENT GROUP

In total, ten patients were investigated for large deletion/duplications in *VWF* (see Table 3.4). These included seven type 3 VWD patients who had been previously identified to have the novel exon 4-5 deletion (c.221-977_532+7059del mutation) discovered in a type 3 VWD study in Manchester. Patients T3-1, T3-3 and T3-4 were previously found to be compound heterozygous for the exon 4-5 deletion and another mutation in exon 8, 37 and 42 respectively. The remaining three patients had been previously found to have either a small deletion or duplication, or a point mutation.

Table 3.4: Type 3 VWD patients investigated by MLPA analysis following screening for mutations within the *VWF*.

PATIENT	M/F	FVIII (U/DL)	VWF: AG (U/DL)	VWF:RCO (IU/DL)	MULTIMERS	RESULT PRIOR TO MLPA ANALYSIS
T3-1	F	2	1	<5		Compound heterozygous Exon 4-5 deletion and c.992_993delGCinsAA mutation in exon 8
T3-2	F	19	4	<5	None detected	Heterozygous for Exon 4-5 deletion
T3-3	F	4	10	12		Compound heterozygous Exon 4-5 deletion and c.6397del14 mutation in exon 37
T3-4	M	13	3	<10	Reduced	Compound heterozygous Exon 4-5 deletion and c.7182T>G mutation in exon 42
T3-5	M	1		<20	None detected	Homozygous for Exon 4-5 deletion
T3-6	M	2	<5	<5	None detected	Homozygous for Exon 4-5 deletion
T3-7	M	2	Not detected	Not detected	None detected	Homozygous for Exon 4-5 deletion
T3-8	F	12	<10	<10	Markedly reduced	Heterozygous for c.5083delCTC mutation in exon 29
T3-9	M	12	0-2	<10	None detected	Compound heterozygous the c.5200C>T and c.7730-3C>G mutations
T3-10	M	3	3	<10		Heterozygous for c.1657dupT mutation

3.4.1 Investigation of Type 3 VWD patients with the exon 4-5 deletion by MLPA

Seven patients were investigated for the presence of the exon 4-5 deletion which had been identified in a type 3 VWD study by Sutherland et al. (2009). In the initial study, patients T3-5, T3-7 and T3-8 (patients T3-7 and T3-8 are siblings) had demonstrated failure to amplify exons 4 and 5 using genomic DNA, which suggested that these exons had undergone a homozygous deletion. Sutherland et al. (2009) had employed the method of 'primer walking' to confirm the deletion and the approximate breakpoints by utilising primer sets which had been designed in overlapping segments. By using these primer sets a PCR product of 1788 bp was produced, rather than the expected product size of 10.4 kb. To be able to map the exact deletion breakpoint by DNA sequencing, a further set of deletion mapping (Delmap) primers were designed which produced a product of 1084 bp rather than the expected 9715 bp in the patients homozygous for the exon 4-5 deletion. A total in-frame deletion of 8631 bp from within intron 3 to a point within intron 5 was identified by sequencing of the 1084 bp product.

Sutherland et al. (2009), also identified four patients (T3-1, T3-2, T3-3 and T3-4) to be heterozygous for the exon 4-5 deletion. In these patients, sequencing of genomic DNA had failed to explain the type 3 VWD phenotype (either no mutation or heterozygosity for only one mutation in *VWF*). Initial PCR analysis of platelet derived *VWF* cDNA produced a truncated PCR product of 683 bp (instead of the expected wild type product of 995 bp) comprising *VWF* exons 2 to 7. Following sequencing analysis of the truncated product, the deletion of the entire coding sequence for exons 4 and 5 was identified (see Figure 3.20). Therefore at the genomic DNA level, the wild type sequence on the other *VWF* allele had masked the presence of the heterozygous deletion during PCR amplification. In all four patients the heterozygous nature of the deletion was confirmed by designing an assay containing two primer sets. One set of primers amplified across the breakpoint (1084 bp product) and the second set amplified across exons 4 and 5 (1694 bp product) as a control. Both products were amplified in all four patients.

Figure 3.20: Example of the identical exon 4-5 deletion breakpoints in patients homozygous and heterozygous for the mutation as identified by Sutherland, et al. (2009).



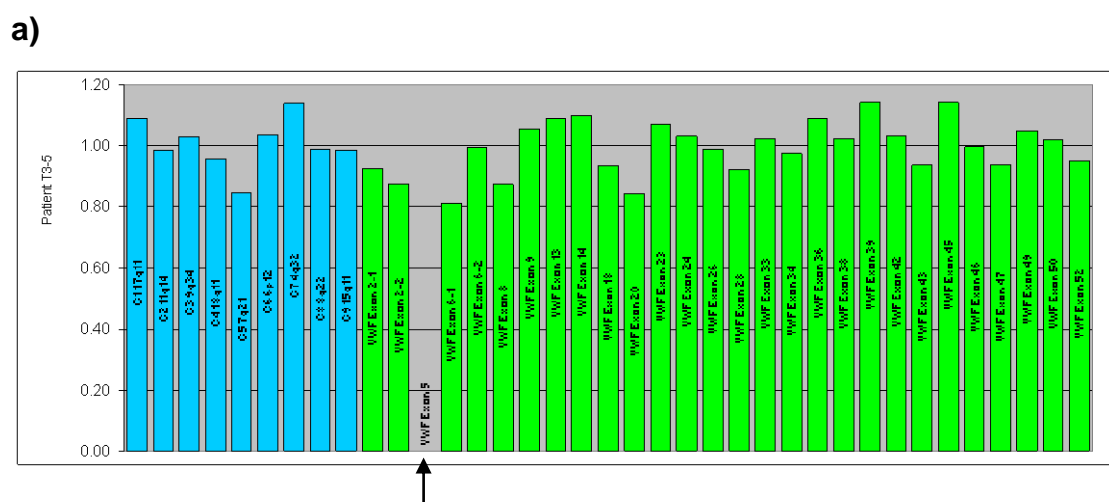
As the presence of the exon 4-5 deletion had been identified by sequencing alone in the Sutherland et al. (2009) study, the MLPA method was used as an extra confirmation as it is specifically applicable to the identification of large deletions/duplications. This is especially pertinent for those mutations of a heterozygous nature which are often masked by the wild type allele during standard PCR amplification.

3.4.2 Confirmation of homozygosity for exon 4-5 deletion by MLPA

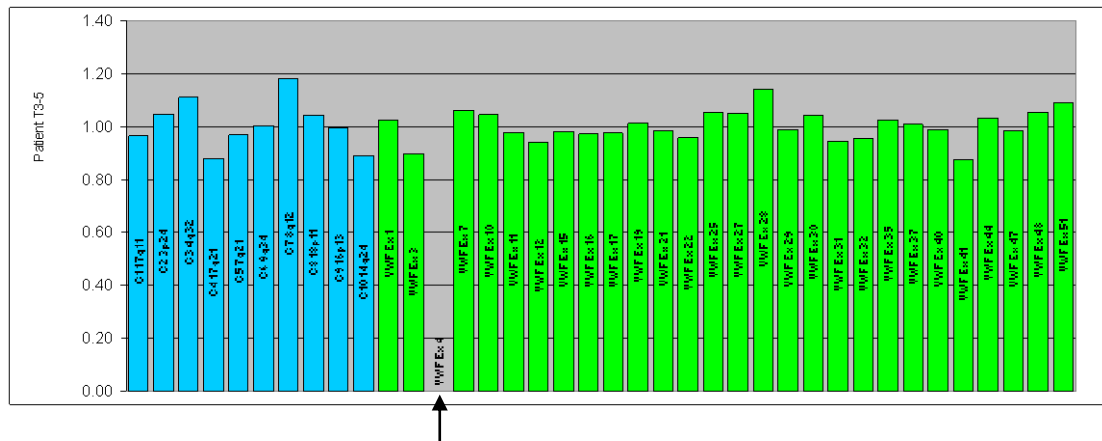
Mutation analysis had been previously performed with an aim to identify the underlying candidate mutation associated with the type 3 VWD in patients T3-5, T3-7 and T3-8. All three were identified to be homozygous for the exon 4-5 deletion of *VWF* which results in a deletion of 8631 bp from the *VWF* sequence. The region that is deleted contains the coding sequence for exons 4 and 5 and is located in the multimerisation domain of VWF. The mutation predicts a deletion of amino acids from aspartate at position 75 to glycine at position 178 inclusive. It is predicted that this mutation will result in either no functional or a truncated VWF protein. Given the nature of the exon 4-5 deletion mutation and the detection of the mutation in all three individuals with type 3 VWD, homozygosity for this mutation is highly likely to account for the type 3 VWD phenotypes seen in these patients.

MLPA analysis carried out on this group of three patients clearly identified the homozygous deletions of exon 4-5 within *VWF*. The analysis was performed at least twice on different occasions for each patient, thereby confirming the initial result. As *VWF* is so large, the MLPA method is split into two kits. The histograms in Figure 3.21 show both kits for *VWF* applied to a DNA sample from patient T3-5. The probes for each exon were amplified apart from exon 5 in kit 1 and exon 4 in kit 2, indicating the presence of a homozygous deletion of exons 4 and 5. Each of the samples for all three patients displayed the same complete absence of amplification of these probes, confirming the previous sequencing results. No evidence of any other large deletions or duplications was identified in any of the samples. The best results were obtained for patient T3-5 with both kits passing quality control. However for T3-7 and T3-8 only kit 1 passed quality control with kit 2 repeatedly failing because of an outlying internal reference probe. Although this increases the possibility of making an incorrect diagnosis, despite this the samples repeatedly showed absence of the exon 4 probe in kit 2, therefore it was clearly obvious that the exon 4-5 deletion was present in these samples.

Figure 3.21: MLPA Histograms for Patient T3-5. The histogram for **a)** Kit 1 shows complete absence of the probe for exon 5 (indicated by the arrow). There is also complete absence of the probe representing exon 4 in histogram **b)** Kit 2 (indicated by the arrow). Together they illustrate the presence of a homozygous deletion of exon 5 and exon 4 of *VWF*.



b)



3.4.3 Confirmation of heterozygosity for exon 4-5 deletion by MLPA

Mutation analysis had been previously performed with an aim to identify the underlying candidate mutation associated with the type 3 VWD in patients T3-1, T3-2 and T3-3 and T3-4. All four were identified to be heterozygous for the exon 4-5 deletion (c.221-977_532+7059delGCinsAA) of *VWF* which results in a deletion of 8631 bp from the *VWF* sequence.

Patient T3-1 had been previously found to have compound heterozygosity for two mutations detected within exons 4-5 and exon 8 of the *VWF* sequence. The second mutation found was the c.992_993delGCinsAA nucleotide change in exon 8 which predicts an insertion-deletion mutation at amino acid 331 causing the deletion of a cysteine amino acid and the insertion of a premature stop codon p.(Cys331delins*). Given the nature of these mutations, compound heterozygosity involving these is highly likely to account for the type 3 VWD phenotype seen in this patient.

Patient T3-2 had been previously diagnosed to have type 3 VWD (FVIII 19 IU/dL, VWF:Ag 4 IU/dL and VWF:RCo <5 IU/dL) however as no other mutation was detected by DNA and RNA analysis in combination with the heterozygous exon 4-5 deletion, a diagnosis of type 3 VWD could not be confirmed by the genotype. Also the reported VWF:Ag level of 4 IU/dL for this patient is more indicative of a diagnosis of severe type 1 VWD rather than type 3 VWD. It is unlikely that heterozygosity for the exon 4-5 deletion alone would account for

the severe phenotype as other heterozygous carriers of this mutation have been seen in mild/moderate type 1 VWD.

Patient T3-3 had been previously found to be compound heterozygous for two mutations with exons 4-5 and exon 37 of the *VWF* sequence. The second mutation identified was the c.6397del14 deletion mutation in exon 37 which predicts a frame shift mutation at amino acid 2133 causing the amino acid valine to be replaced by proline and the insertion of a premature stop codon 12 codons downstream p.(Val2133Profs*12). It is predicted that this mutation will result in a null allele and so no mature VWF protein will be produced. Given the nature of the exon 4-5 deletion mutation and c.6397del14 mutations, compound heterozygosity for these mutations is highly likely to account for the severe phenotype seen in this patient.

Patient T3-4 had also been previously found to be compound heterozygous for the exon 4-5 mutation and a c.7182T>G nucleotide change in exon 42 which predicts a missense mutation at amino acid 2394 p.(Cys2394Trp). It is predicted that this removal of the cysteine amino acid may cause disruption to the tertiary structure of the mature VWF protein resulting in a dysfunctional protein. In this respect, it is worth noting that patient T3-4 has a reported VWF:Ag level of 3 IU/dL and a diagnosis of severe type 1 VWD rather than type 3 VWD cannot be excluded. Given the nature of the exon 4-5 deletion mutation and c.7182T>G mutations, compound heterozygosity for these mutations is highly likely to account for the severe type 1 VWD phenotype seen in this patient.

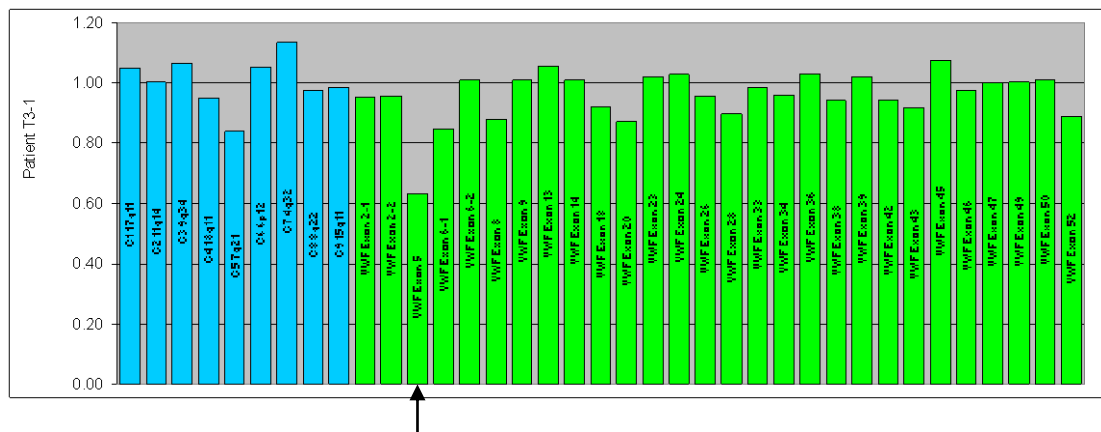
MLPA analysis carried out on this group of four patients clearly revealed the heterozygous deletions of exon 4-5 within *VWF*. Patient T3-1 was the only patient from this group who displayed good quality results for both kits (standard deviation of <0.1). The histograms in Figure 3.22 show both kits for *VWF* applied to a DNA sample from this patient. It can be seen that the probe heights for exon 5 in kit 1 and exon 4 in kit 2 are significantly reduced and when measured against the internal reference probes are half the height, indicating the presence of a heterozygous deletion of exons 4 and 5. This was also

observed in the other three patients, however at least one kit failed quality control on numerous occasions for each of these patients. With respect to patients T3-1, T3-3 and T3-4 who were shown to have compound heterozygosity by previous mutation analysis, this was not evident by MLPA. The insertion-deletion mutation (c.992_993delGCinsAA) in exon 8 identified in patient T3-1 did not have an effect on the peak height of the probe for exon 8. The small deletion mutation (c.6397del14) found in patient T3-3 which results in the deletion of 14 bases from exon 37 also was not found to have an effect on the peak height of the probe for exon 37 and therefore provided no evidence to the presence of this mutation. This was also the same for patient T3-4 who is heterozygous for the missense mutation c.7182T>G in exon 42 which results in a single amino acid substitution in the mature VWF protein. No evidence of a reduced peak height was seen for exon 42, which would indicate the presence of this mutation.

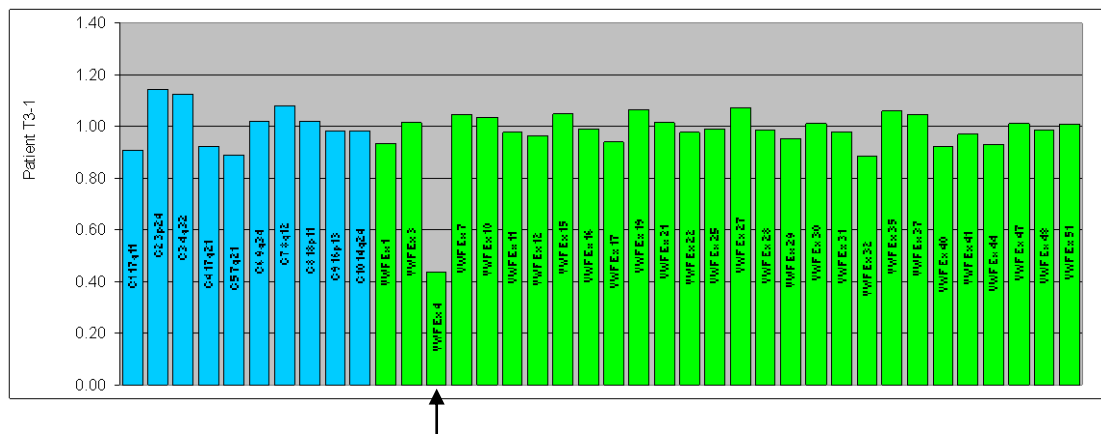
Although only 1 out of the 4 patient samples passed quality control, together with the initial results from mutation analysis the presence of the heterozygous exon 4-5 deletion was clearly identifiable. Following this confirmation in a diagnostic laboratory setting, this information could be used in future genetic family studies as required.

Figure 3.22: MLPA Histograms for Patient T3-1. The histogram for **a)** Kit 1 shows a reduced peak height of the probe for exon 5 (indicated by the arrow). There is also a reduced peak height of the probe representing exon 4 in histogram **b)** Kit 2 (indicated by the arrow). Together they illustrate the presence of a heterozygous deletion of exon 5 and exon 4 of *VWF*.

a)



b)



3.4.4 Investigation of type 3 patients with no causative mutation by MLPA

The three remaining patients in the type 3 VWD group to be investigated had also previously had mutation analysis of *VWF* carried out in order to identify the underlying candidate mutation associated with their type 3 VWD diagnoses.

Patient T3-8 had previously been found to be heterozygous for a previously unreported c.5083delCTC deletion mutation within exon 29 which results in a deletion of 3 base pairs from the *VWF* sequence. As a result the amino acid leucine at position 1695 is deleted. It is predicted that this mutation may result

in an abnormal tertiary protein structure which may reduce the function of the mature VWF protein. No other mutation was detected by DNA or RNA analysis in combination with this mutation and therefore the diagnosis of type 3 VWD could not be confirmed by the genotype. As patient T3-8 had a reported VWF:Ag level of <10 IU/dL with detectable VWF multimers, a diagnosis of severe type 1 VWD rather than type 3 VWD may be more likely. It is not clear how heterozygosity for the identified p.Leu1695del mutation would alone produce a severe VWD phenotype.

Patient T3-9 was previously found to show compound heterozygosity for two DNA single base changes. The c.5200C>T base change within exon 30 was identified which predicts the nonsense mutation glutamine 1734X, or p.(Gln1734X), of the mature VWF protein. The second base change identified was the c.7730-3C>G base change which is likely to affect a splice site at the intron 45, exon 46 boundary. Although neither of these *VWF* mutations have been previously reported, given their nature they are highly likely to account for the severe VWD phenotype seen in this patient.

Patient T3-10 was previously diagnosed to have haemophilia A, however more recent phenotypic laboratory tests indicated a revised diagnosis of severe VWD. Previous mutation analysis of *VWF* had identified heterozygosity for a DNA single base duplication within exon 14 of the *VWF* sequence (c.1657dupT). This nucleotide duplication predicts the amino acid substitution tryptophan 553 to leucine, resulting in a frame shift mutation which introduces a premature stop signal 97 codons downstream p.(Trp553Leufs*97) in the VWF protein. There is a previous report of this mutation on the ISTH *VWF* database in association with type 3 VWD when inherited in compound heterozygosity with a second *VWF* mutation. It is unlikely that heterozygosity for this mutation would alone account for the severe VWD phenotype seen in this patient. A second *VWF* mutation was not identified, however it cannot be excluded that an additional mutation is present that cannot be detected using conventional DNA sequencing methodology.

MLPA analysis was performed on these three patients to try and identify whether a large deletion may be present which would account for their severe phenotype. Both patients T3-9 and T3-10 showed good quality results for both kits (standard deviation of <0.1), however for patient T3-8 *VWF* kit 2 was just over the 0.1 threshold. This was probably due to the fact that the DNA sample for patient T3-8 was much older than the DNA used for analysis for the other two patients. However it was still possible to obtain a clear picture of any mutations that may be responsible for the severe type 3 phenotype. Both T3-8 and T3-10 showed normal peak heights for all 52 exons within *VWF*. Although patient T3-8 had previously been identified to have a small deletion of 3 bp from exon 29 this was not identified by the MLPA kit. A similar result was found for patient T3-10 who had previously been identified to have a single base duplication within exon 14, which was not detected by the MLPA kit.

As described earlier, patient T3-9 had previously been found to be compound heterozygous for two different base changes. MLPA analysis showed a reduced peak height for the probe for exon 30 which coincided with the presence of the c.5200C>T base change within exon 30 that had been previously detected. This is illustrated in Figure 3.23 which shows the results for patient T3-9 using *VWF* kit 2 and the drop out of exon 30. There was no evidence in kit 1 of the second base change (c.7730-3C>G) affecting the peak heights of the probes for either 45 or 46. It is well known that the presence of a mutation or polymorphism in the sequence detected by a probe can cause a reduction in the relative peak height which had also been seen in a haemophilia family using the *F8* kit (see section 3.3.2.10). Figure 3.24 shows the gene transcript for exon 30 and the location of the *VWF* MLPA probe highlighted in pink and green. The location of the c.5200C>T base change is highlighted in purple which is present within the probe sequence. Therefore if this base change is present it will cause the reduction in peak height of a probe, indicating a potential limitation of the MLPA method.

Figure 3.23: MLPA Histogram for Patient T3-9. The histogram for VWF Kit 2 shows a reduced peak height of the probe for exon 30 (indicated by the arrow) caused by the c.5200C>T base change within exon 30 in this patient.

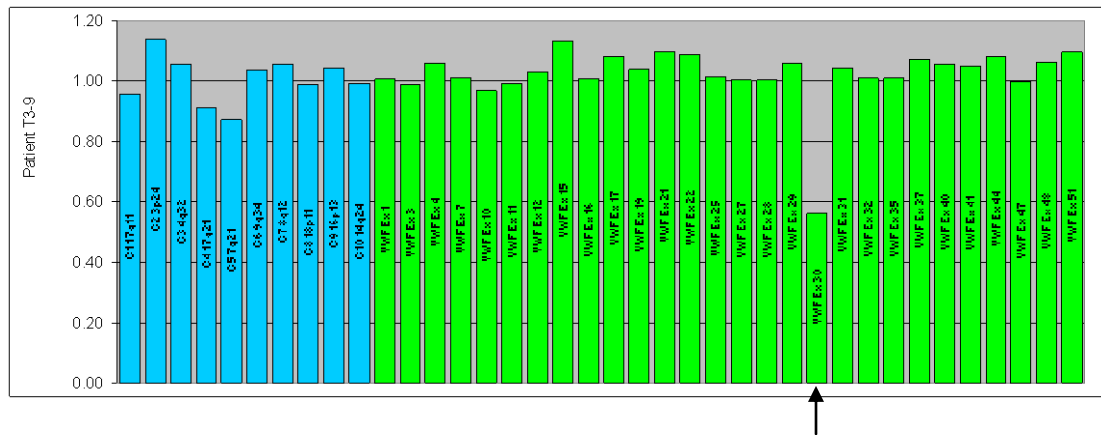


Figure 3.24: VWF transcript for exon 30. The transcript shows the location of the exon 30 probe (highlighted in pink and green) and the location of the c.5200C>T base change (highlighted in purple) which leads to a reduction in height of the probe.

```

5401  GCTTTTCATTTCMCAAAAGCCAATATAGGG30YCCTCGTCTCACTCAGGTGTCAGTGCTGCAGTAT
5146  GCTTTTCATTTCMCAAAAGCCAATATAGGGCCTCGTCTCACTCAGGTGTCAGTGCTGCAGTAT
1716  -A--F--I--S--K--A--N--I--G--P--R--L--T--Q--V--S--V--L--Q--Y--

5461  GGAAGCATCACCACCATTGACGTGCCATGGAACGTGGTCCCGGAGAAAGCCCATTTGCTG
5206  GGAAGCATCACCACCATTGACGTGCCATGGAACGTGGTCCCGGAGAAAGCCCATTTGCTG
1736  -G--S--I--T--T--I--D--V--P--W--N--V--V--P--E--K--A--H--L--L--

5521  AGCCTTGTGGACGTCATGCAGCGGGAGGGAGGCCCCAGCCAAATCGGGGATGCCTTGGGC
5266  AGCCTTGTGGACGTCATGCAGCGGGAGGGAGGGAGGCCCCAGCCAAATCGGGGATGCCTTGGGC
1756  -S--L--V--D--V--M--Q--R--E--G--G--P--S--Q--I--G--D--A--L--G--

```

3.5 RESULTS OF MLPA ANALYSIS IN THE TYPE 1 VWD PATIENT GROUP

In total, sixteen patients were investigated for large deletion/duplications in VWF who had originally been recruited into a UK national study (Cumming, Grundy et al. 2006). In the original study, genomic DNA from thirty-two index cases and their families was analysed for mutations in VWF which may account for the type 1 VWD bleeding phenotype. Mutations had not been found in all

patients and families studied therefore RNA work was undertaken by Sutherland et al. (2009) to search for mutations which could not be detected using the conventional DNA PCR techniques. This also included the aim of investigating these patients for the presence of the large heterozygous exon 4-5 deletion which had been detected at a high frequency in the type 3 VWD patient cohort. Of the initial 32 index cases, only 16 had DNA samples available which could be used for MLPA analysis. Phenotypic data for the sixteen patients in whom MLPA analysis was completed are shown in table 3.5. Of the sixteen cases, two index type 1 VWD patients were included who had been previously identified by Sutherland et al. (2009) to be heterozygous for the novel exon 4-5 deletion (c.221-977_532+7059del mutation). The remaining fourteen patients had been previously found to either have no mutation found or putative for mutations which did not co-segregate with their VWD phenotype.

Table 3.5: Type 1 VWD index patients investigated by MLPA analysis following screening for mutations within VWF. The phenotypic laboratory data were provided by the referring centre and by the central testing laboratory.

PATIENT	M/F	VWF: AG (U/DL)	VWF:RCO (IU/DL)	VWF:CB (IU/DL)	BLOOD GROUP	RESULT PRIOR TO MLPA ANALYSIS
T1-2	M	30, 22, 76	33, 37 (70)	(77)	O	None
T1-3	F	13, 16, 13, 21	11, 10, 18, 10, 13	12, 9	O	p.Tyr1584Cys Ex 28 Heterozygous for Exon 4-5 deletion
T1-5	F	33, 51, 59, 69	18, 54, 56, 87	176	O	None
T1-8	F	35, 76	41, 63	30, 106	A	None
T1-10	M	30, 27, 40	23, 24, 22	46, 55	A	p.Arg924Gln Ex 21 c.5843-8C>G Int 34
T1-12	F	27, 31, 37, 46	33 (58)	35,27,68 (49)	A	None
T1-14	F	81, 54, 60, 66, 48	41, 56, 71, 50 (43)	(84)	O	p.Gly2705Arg
T1-15	F	31, 44	25 (56)	(36)	O	p.Arg924Gln Ex 21 p.Val1850Met Ex 32 c.5834-8C>G Int 34
T1-16	F	36, 38	14 (44)	(50)	O	p.Tyr1584Cys Ex 28
T1-18	F	34, 19, 66	32, 25, 72	86	O	c.5170+10C>T Int 29
T1-19	M	48, 61, 57, 51	60	23, 59, 41, 73	O	c.Tyr1584Cys Ex 28
T1-20	F	46, 87	33, 78	106	O	None
T1-23	F	20, 27, 13, 26	115, 53, 28, 87, 22, <10	3	O	Heterozygous for Exon 4-5 deletion
T1-25	F	45, 61, 82, 44, 66, 67, 30, 74	44, 68, 82, 44, 68, 110, 54, 62	39	O	p.Tyr1584Cys Ex 28
T1-28	F	36, 38, 40	36, 39, 42	38	O	p.Asp141Gly Ex 5
T1-30	M	39, 36, 36, 22, 34, 40, 49, 73, 59, 72, 63, 82, 61	28, 23, 42, 16, 21, 21, 41, 61, 70, 55, 59, 92, 63	Not tested	O	p.Thr2647Met

3.5.1 Investigation of Type 1 VWD index patients with exon 4-5 deletion by MLPA

In order to corroborate the work by Sutherland et al. (2009), two index cases (T1-3 and T1-23) that had been identified to be heterozygous for the exon 4-5 deletion were investigated by MLPA analysis. They had previously been analysed for the exon 4-5 deletion using the Delmap primers at the genomic DNA level, with confirmation by a multiplex assay. No candidate *VWF* mutation had been previously identified in the study carried out by Cumming, Grundy et al. (2006). Sequencing of the exon 4-5 deletion PCR products from the Delmap primers confirmed that the deletion carried exactly the same breakpoints that had been identified in the type 3 VWD patients with the mutation (c.221-977_532+7059del). Previous haplotype analysis carried out in family members in both families indicated that the mutation segregated with the same common haplotype and therefore was causative of the type 1 VWD. The *VWF* polymorphism p.Tyr1584Cys also found in patient T1-3 did not segregate with the disease phenotype.

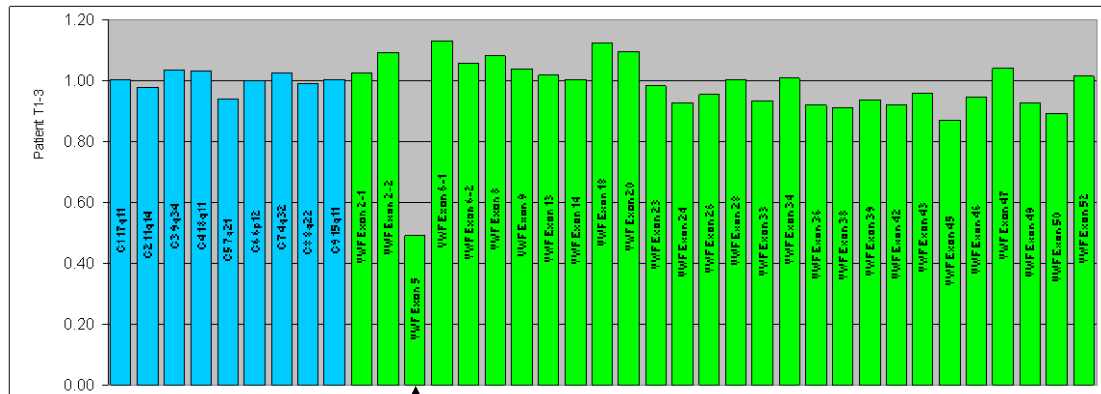
MLPA analysis carried out on these two index patients clearly revealed the heterozygous deletions of exon 4-5 within *VWF*. Patients T1-3 and T1-23 displayed good quality results for both *VWF* kits (standard deviation of <0.1). The histograms in Figure 3.25 show both kits for *VWF* applied to a DNA sample from patient T1-3. It can be seen that the probe heights for exon 5 in kit 1 and exon 4 in kit 2 are significantly reduced and when measured against the internal reference probes are half the height, indicating the presence of a heterozygous deletion of exons 4 and 5. The same was observed in the DNA sample for patient T1-23.

The *VWF* polymorphism p.Tyr1584Cys found in patient T1-3 which is a missense mutation c.4751A>G in exon 28 was not found to have an effect on the peak height of the probe for exon 28, and therefore its presence was not evident by MLPA.

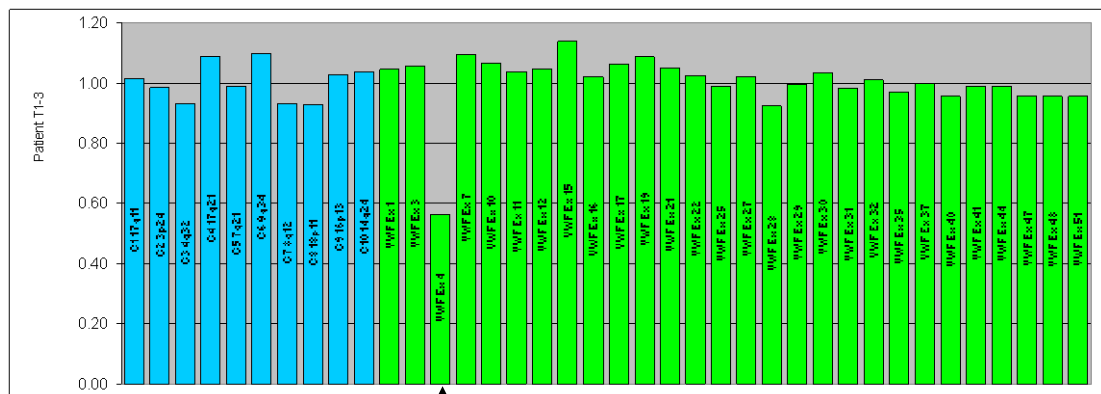
Following this confirmation in a diagnostic laboratory setting, this information could be used in future genetic family studies as required.

Figure 3.25: MLPA Histograms for Patient T1-3. The histogram for **a)** Kit 1 shows a reduced peak height of the probe for exon 5 (indicated by the arrow). There is also a reduced peak height of the probe representing exon 4 in histogram **b)** Kit 2 (indicated by the arrow). Together they illustrate the presence of a heterozygous deletion of exon 5 and exon 4 of *VWF*.

a)



b)



3.5.2 Investigation of type 1 VWD index patients with no causative mutation by MLPA

The fourteen remaining index patients in the type 1 VWD group to be investigated had previously had mutation analysis of *VWF* carried out in order to identify the underlying candidate mutation associated with their type 1 VWD diagnoses. These patients were either found to have no putative mutations or the identified putative mutation did not co-segregate with the VWD phenotype. Therefore in an attempt to ascertain whether a large deletion may be present

which may be responsible for the type 1 VWD phenotype, MLPA analysis was carried out.

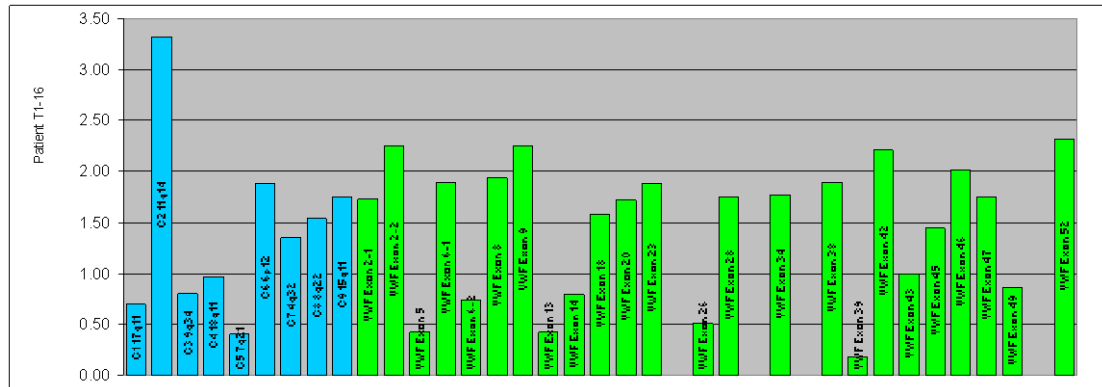
MLPA analysis carried out on this group of type 1 VWD patients failed to identify any large deletions within *VWF*. In this group there was a high failure rate with respect to samples passing quality control. Even though analysis was performed at least three times for each kit, samples repeatedly failed. Only four out of the fourteen patient samples analysed passed quality control for both kits 1 and 2 for *VWF*. These four patients (T1-2, T1-8, T1-14, T1-19) produced normal results with no evidence of large deletions which may be responsible for their phenotype.

Three of the fourteen patients (T1-10, T1-15 and T1-16) completely failed quality control for both kits and therefore could not be analysed. On many occasions probes failed to ligate to target DNA sequences which was made apparent by the complete lack of probe amplification. Figure 3.26 shows the histograms for patient T1-16 where for both kits the internal reference probes failed together with inadequate amplification of probes for each exon. Figure 3.26a is an example of how this result could lead to an incorrect diagnosis of a large deletion if the standard deviation is not taken into account as several of the probes for individual exons have failed to amplify. When MLPA analysis was repeated on this sample, all the probes were amplified although the sample still failed quality control. The remaining seven patients could only partially be analysed as only kit 1 passed quality control and kit 2 repeatedly failed, therefore a complete diagnosis by MLPA was not obtained for these patients.

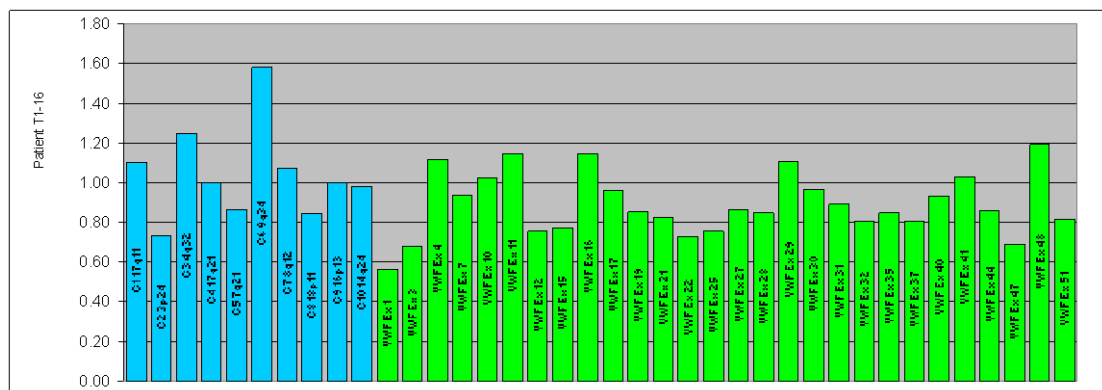
It was difficult to ascertain whether some of the putative *VWF* polymorphisms identified by DNA sequencing in some of the patients had an effect on probe height. This was due to the fact the majority of the exons in which they were present were part of the *VWF* kit 2 which consistently failed for many patients.

Figure 3.26: MLPA Histograms for Patient T1-16. The histograms for both **a)** Kit 1 and **b)** Kit 2 illustrate irregular probe amplification for the internal reference probes (blue bars) and individual exons (green bars). This results in sample failure as the standard deviation will be >0.1 .

a)



b)



CHAPTER 4: DISCUSSION

4.1 INTRODUCTION

Haemophilia A is caused by a broad spectrum of mutations which are spread throughout the 26 exons of *F8*. These mutations can be identified through mutation screening which facilitates carrier detection in female family members and can be helpful in determining the risk of complications such as inhibitor formation in affected males. Several techniques are used for mutation screening including long-range and/or inverse PCR for detection of the inversions of introns 1 and 22 which account for 45-50% of severe haemophilia A patients, and sequencing of the coding regions of *F8* to identify point mutations, deletions or splice-site mutations in the majority of remaining cases. However a small proportion of haemophilia patients do not have a mutation that can be identified using these techniques. Duplications of part of *F8* are not easily detected by these means as gene dosage is not established. It is also difficult to determine carrier status in female relatives of males with large *F8* deletions using simple targeted PCR as they are masked by amplification of the normal X chromosome.

The diagnosis of VWD is primarily based on clinical and phenotypic information, however more recently genetic analysis has become available to confirm or support the initial diagnosis and provide information for family studies and prenatal diagnosis which is particularly important in type 3 VWD. Genetic analysis is useful in the differential diagnosis of VWD type 2N and mild haemophilia A, or VWD type 2B and platelet type pseudo-VWD. Genetic investigation is not routinely used in type 1 VWD as it has been shown to be complicated due to a causative mutation not always being detectable (Cumming, Grundy et al. 2006; Goodeve 2007; James, Notley et al. 2007). A large range of mutations have been reported to be associated with VWD including missense, nonsense and splice site mutations, small deletions/insertions and large deletions. Large deletions responsible for VWD are relatively uncommon, however they usually result in a complete deficiency of VWF protein (Schneppenheim, Castaman et al. 2007). To date, twelve large deletions have been reported for type 3 VWD, ranging in size from a single

exon to the entire gene. Large deletions are difficult to detect in autosomal conditions as amplification of the normal allele masks the failure of the other allele to amplify. Therefore, false-negative results may be obtained for carriers of heterozygous large deletions.

MLPA is a method which has the potential to overcome the above limitations. It has recently been employed to identify large deletions and duplications within *F8* and *VWF* (Rost, Loffler et al. 2008; Lannoy, Abinet et al. 2009; Zimmermann, Oldenburg et al. 2010; Cabrera, Casana et al. 2011; Rafati, Ravanbod et al. 2011; Theophilus, Baugh et al. 2011; Yadegari, Driesen et al. 2011). The method is used to screen for the presence of deletions or duplications by quantitatively comparing the copy numbers of a set of DNA sequences in a patient sample to those in a control sample.

The aim of this study was to apply MLPA to investigate the molecular basis of disease in patients with haemophilia A and VWD where no causative mutation had been identified using the 'gold standard' method of direct sequencing. In order to do this, it was essential to first develop, optimise and validate the MLPA assays for *F8* and *VWF* in order to identify large deletions/duplications. The methodology was then applied to a selected cohort of patients with Haemophilia A and VWD to further investigate the underlying cause of their disease.

4.2 MLPA METHOD EVALUATION

When MLPA was selected as the method to detect large deletions and duplications it was important to take into account a number of considerations before starting such as available resources, ease of method implementation in the laboratory, experimental set-up, sample treatment and interpretation of results.

The necessary equipment for MLPA consists of a thermocycler and capillary sequence equipment which was already available in the laboratory, therefore the initial investments costs were low. An important part of the initial experimental set-up was to test the thermocycler to be used for excessive

evaporation as this can affect MLPA results. Excessive evaporation can be caused by a particular type of tube, the pressure of the heated lid on the tube or whether the heated lid is working properly. The plastic of the tubes can be weak at high temperatures and lead to deformation of the tubes when the pressure of the lid is too high which results in tube leaks and excessive evaporation. A test for excessive evaporation was carried out on the thermocycler using two different brands of tubes. One brand of tube was found to have less evaporation than the other and was used in all further experiments. The other brand had warped overnight in the thermocycler and caused more evaporation which was probably due to a weaker plastic used in tube manufacture.

MLPA kits contain everything that is required for an assay (probemix, buffers, ligase, polymerase, dNTP's and labelled PCR primers) without the need to buy additional reagents. An advantage of MLPA is that it can be used for many different applications and the only difference between MLPA assays is the probemix specific to each target region of DNA. This was of particular value for this study as the method could be evaluated for both Haemophilia A and VWD at the same time with the same conditions.

An important factor that was taken into account was the sensitivity of the MLPA method as this can often be a limitation. MLPA literature states that it is a very sensitive technique where only 20 ng of human DNA is required in an assay and results do not vary depending on the amount of sample DNA used. It can distinguish sequences differing in only a single nucleotide, if a specific probe has been designed for this purpose and is able to detect small copy number variations in a DNA sample. It can also detect deletions or duplications of a single exon as it detects sequences of only ~60 nt (evidence of this is described in section 4.3.1).

MLPA assays for *F8* and *VWF* were first performed on control DNA samples from healthy individuals which provided good initial results. It is necessary to include reference samples in each experiment as minor differences between different experiments may affect the MLPA peak pattern. As recommended the

control samples were derived from the same type of tissue and extracted using the same protocol as the test samples. Although it is recommended that at least five control samples should be included in each run, it was decided that seven controls should be used. This proved to be advantageous as having two extra controls covered for any inter assay control failures which may occur, meaning that the whole run would not have to be repeated. It is not essential to include positive controls, therefore this was not incorporated into the final protocol for MLPA.

4.2.1 Sample treatment

The starting material for analysis was of particular importance for MLPA. The treatment of DNA (storage and extraction method) may affect MLPA results. Although special DNA extraction methods are not required for MLPA analysis, the extraction method should not leave behind a high concentration of contaminants such as salt. A high salt concentration can compromise MLPA as it prevents complete denaturation of the CG rich parts of the sample DNA (CpG) islands. Incomplete denaturation can lead to false positive MLPA results, including an apparent deletion of several probes. It is also not recommended to use heparinised blood as traces of heparin can still be present in DNA preparations which can lead to distortion of the MLPA PCR reaction. Fortunately all DNA samples used in the study had been previously extracted from peripheral blood samples in EDTA by an ammonium acetate method, therefore this did not pose an initial problem. However it is also recommended that for best results the DNA should be extracted from fresh samples and that a concentration of between 50-250 ng of DNA in 5 μ l should be used for each MLPA reaction. As the samples used in this study were from patients who had already had previous genetic studies carried out, they were obtained from stocks that had been frozen for a number of years. Therefore DNA degradation had occurred and the concentration was often low. This led to consistently poor results for some patients with very old samples where only a fresh sample would give results that could be reliably reported. Impurities in the extracted DNA which have an effect on MLPA results can also be lowered by reducing the amount of sample DNA, therefore any impure samples were diluted to a working stock of 10-20 ng/ μ l. This was found to be beneficial in a few

circumstances but generally led to ligation and amplification failure. RNA contamination of the DNA samples can also affect the results of a small number of probes that detect sequences in genes that are highly expressed e.g. globin α and β . This is due to RNA molecules reducing the effective probe concentration by binding to the probe or by competing with the probes for binding to the same target sequence on the genomic DNA. The amount of RNA also differs from sample to sample, which results in a high variability of certain probes. Although *F8* and *VWF* are not transcribed in abundance, it was necessary to test whether the addition of an RNase treatment would be of benefit to those samples with consistently poor results. Similarly to reducing the DNA concentration to reduce impurities, the addition of the RNase treatment was advantageous for some failed samples but not all.

Following these attempts to optimise sample treatment it was concluded that the best results were obtained from DNA extracted from fresh peripheral blood samples in EDTA. In the general setting of a routine genetic service, MLPA analysis will potentially be included as part of the initial diagnostic algorithm and therefore fresh samples will be analysed. However in cases where retrospective analysis needs to be carried out on older samples such as in this study, this may pose a problem in generating reliable results for reporting. Lowering the concentration of an impure stock DNA sample did not have any major beneficial effect and if possible a new sample should be obtained. However the implementation of the RNase treatment as a standard part of the sample treatment is deemed to be useful as it can only have a positive effect and not a negative one.

4.2.2 Method Optimisation

Attempts were made to optimise the MLPA protocol in order to achieve the most consistent results. A standard protocol is supplied by MRC-Holland but slight changes to this may result in minor improvements to data quality. Initial experiments were all tested according to the standard protocol, however for those samples that consistently failed quality control it was decided that a slightly different protocol optimised by Edinburgh Royal Infirmary should be tested to establish whether better quality results could be obtained. The main

difference between the two protocols was the number of PCR cycles for the PCR reaction. The standard protocol uses 35 PCR cycles however the Edinburgh protocol had reduced the number of cycles to 33 which can slightly improve the linearity of relative probe signal with target sequence copy number. Samples were tested in parallel using the two protocols and data quality proved to be better using the standard protocol with better peak amplification and morphology. Therefore it was decided that the standard protocol would be adopted as the routine method for all future samples.

4.2.3 Data analysis

The correct evaluation of data is crucial for the reliable implementation of MLPA analysis. Firstly, the peaks generated by fragment analysis need to be visually evaluated in order to prevent false positive results. Based upon this evaluation the MLPA procedure may need to be repeated, e.g. samples may need to be tested again on the capillary sequencer when peaks are either above the fluorescence maximum of the sequencer or when they are very low. Other problems may be identified through visual examination of the peak patterns of the size-called data, including the size standard, such as the use of deteriorated capillary electrophoresis gel matrix, old capillaries or low quality formamide. It is also important to assess the control fragments with amplification products between 64 nt and 105 nt and the 'no DNA control' reaction. The four DNA quantity control fragments (Q-fragments) and one ligation-dependant control fragment should be evaluated to ensure that sufficient sample DNA was present and that the ligation reaction was successful. The two DNA denaturation fragments (D-fragments) should also be assessed for incomplete denaturation together with the X-chromosome specific and Y-chromosome specific control fragments. The 'no DNA control' reaction should only generate a probe signal for the Q-fragments as they do not require a ligation reaction or the presence of sample DNA to be amplified. However non-specific peaks may be seen in the 'no DNA control' due to the MLPA PCR reaction which includes a high number of long MLPA probe oligonucleotides that each contain an exact copy of one of the PCR primers. However, in most cases, the non-specific peaks do not have an effect on the MLPA results of

normal samples as they are outcompeted by genuine MLPA probe amplification products.

Non-specific fragments also occur in the region of the control fragments. A non-specific fragment was often seen at ~88 nt. This was due to the use of the ABI POP7 gel matrix for capillary electrophoresis and it could be ignored as it represents non-specific amplification of a long primer-dimer that has the same mobility as one of the quality fragments. If the non-specific peak affected the analysis of the 88 nt D-fragment then only the 96 nt D-fragment peak was used.

4.2.4 Result interpretation

Data for both *F8* and *VWF* were normalised using spreadsheets obtained from the NGRL Manchester. In order to verify whether the calculated ratios are reliable, it is essential to be familiar with the MLPA technique and the gene it is being applied to. When interpreting the data a number of considerations need to be taken into account before a final conclusion can be made. False positive results can be produced by sequence changes including single nucleotide polymorphisms (SNPs) or point mutations in the target sequence detected by a probe. The probe signal can be inhibited by mutations/SNPs in close proximity to the ligation site (1-5 nt) as they can prevent ligation of the two probe oligonucleotides. Mutations/SNPs that are further away from the probe ligation site (at least 20 nt) can produce a lower probe signal by unbalancing the binding of the probe oligonucleotide to the sample DNA. Although the assays have been designed to avoid SNPs if possible, the increasing number of described SNPs shows it is likely that in some individuals rare SNPs may interfere with probe binding. With this in mind, confirmation by another method is always required for copy number changes that are detected by a single probe only. Sequencing is usually the method of choice for determining whether there is a polymorphism or mutation present in the probe target sequence whereas long range PCR and qPCR are often used to determine exon deletions (see section 4.3). The frequency of single exon deletions is generally higher in genes with large introns than in small genes with very short introns which was worth noting in this study as both *F8* and *VWF* are large genes. The pathogenic nature of any deletions or duplications detected by MLPA also

needs to be determined as not all result in disease. Partial gene duplications within the gene may disrupt that copy of the gene resulting in a pathogenic outcome, while a duplication of the complete gene at a different chromosomal location might not. In the majority of diseases the underlying genetic defect is due to a small point mutation which will most probably not be detected using the MLPA approach, unless a probe has been designed specifically to detect these. Also MLPA will not pick up most inversions, translocations or copy number changes which are not within the sequence detected by the probes. This was an important factor to consider when using MLPA to investigate haemophilia A and VWD and section 4.5 describes how MLPA has been incorporated into the mutation detection strategies for both diseases.

4.2.5 Advantages and Limitations of MLPA

The application of MLPA for copy number detection offers many advantages over other techniques. Methods such as sequencing which were mainly developed for detection of point mutations, small deletions and insertions, usually fail to detect copy number changes. Depending on the size, not all deletions can be amplified by sequencing if the primers from neighbouring exons are too far apart, particularly if the deletion has not been mapped. This is why MLPA offers an advantage as an indirect screening assay in a diagnostic context when the presence of a deletion or duplication has not yet been identified. It offers a faster alternative to developing bespoke direct assays for the detection of deletions and it is a standardised, robust method covering an entire gene or locus. By introducing MLPA to initially screen for a deletion/duplication a direct assay can then be designed when the location of the deletion/duplication has been determined, if required. However, although the approach of designing a long PCR reaction followed by sequencing may detect fewer deletions than MLPA, it is the 'gold standard' method as it gives direct confirmation of the nature of the deletion. In this study, MLPA was used to support the identification of large deletions/duplications initially discovered through direct assays, for example the novel exon 4-5 deletion (c.221-977_532+7059del mutation) found in type 1 and type 3 VWD patients.

MLPA is used mainly to detect genomic deletions/duplications and is not suitable for identifying unknown point mutations. However probes can be designed so that the ligation site is located exactly at the position of a known (point) mutation. Therefore, ligation will only occur when the mutation is not present and in the case of a mutation will result in a decreased fluorescent signal.

As discussed in section 4.2.4, MLPA is sensitive to small deletions, insertions and mismatches which could affect the probe signal. Therefore a non-pathogenic polymorphism could be mistaken for a deletion due to a decreased probe signal. The latest databases are used to avoid SNPs when designing MLPA probes, however it is impossible to exclude all cases. This is why MLPA findings should be confirmed by another method. In this study, a reduced probe signal (caused by mutations in the sequence detected by the MLPA probe) was observed in two haemophilia cases (Patients 18 and 19) and one type 3 VWD case (patient T3-9). Patient 18 was a severe haemophiliac who had previously been found to have a small 14 bp duplication within exon 17 of *F8*. His sister (Patient 19) had also been investigated and was found to be a carrier of the same small duplication mutation. When MLPA analysis was carried out on samples from these patients a reduction in the relative peak height was clearly seen for exon 17 due to the presence of the mutation within the sequence detected by the probe. In the case of patient T3-9, previous mutation analysis had identified compound heterozygosity for two base changes. One of the base changes within exon 30 of *VWF* (c.5200C>T) caused a reduction in the peak height for exon 30 due to the presence of the mutation within the sequence detected by the probe. In these three cases, the reduction of the peak height could be misinterpreted as a deletion, however the presence of these mutations had already been confirmed by other methods before MLPA was carried out. In a normal routine diagnostic setting, indirect screening for deletions/duplications using MLPA would initially be carried out followed by the design of a direct assay.

Although MLPA has proved to be a sensitive and robust method when performed to a high standard, it does have some limitations. Compared to

conventional PCR, MLPA reactions are more sensitive to contaminants such as phenol, ethanol, heparin and guanidinium salts. However the presence of contaminants will be readily detected as strongly reduced probe signals will be seen for the longer probes. To overcome this problem it is possible to clean contaminated samples by ethanol precipitation or use smaller amounts of DNA in the MLPA reaction to reduce the number of impurities.

Currently MLPA probe sets are available from MRC-Holland and to develop in-house probe sets would be complicated, expensive and time-consuming. Each probe requires primer design and preparation of a phage M13 clone, purification of its single-stranded DNA and subsequent digestion with expensive restriction endonucleases. However it is possible to send in requests to MRC-Holland for new probe mixes.

MLPA does not have the ability to analyse single cells, such as in the case of FISH. MLPA analysis will give the average copy number per cell in a DNA sample with a mixed cell population. Therefore if DNA was extracted from a sample that contained less than 50% cancer cells, it would be difficult to detect deletions of a certain gene, which would be the case in tumour analysis.

4.3 HAEMOPHILIA A PATIENTS

In the present study, large deletions were confirmed by MLPA analysis in seven of the fourteen index cases investigated. All of the large deletions had been previously identified through routine screening, however not all of the cases had direct assays designed to determine the breakpoints of the deletions identified. MLPA was also applied to the definition of carrier status in five female patients (four were relatives of the index cases investigated). Two females (Patients 8 and 13) were confirmed to be carriers of the large deletion found in the index case, while the remaining three (Patients 6, 10, 19) were found to be negative for any large deletions/duplications. One severe female haemophiliac (Patient 16) was found to have a possible combined duplication and deletion event whereas her severe haemophiliac nephew (Patient 17) was found to be mutation-negative, however, there were issues with the sample quality in these cases as they were old (see section 4.3.1.3). A small

duplication mutation previously found in one severe index case (Patient 18) and his sister (Patient 19) showed how this caused dropout of the MLPA probe signal for one exon (see section 4.2.5). All four of the mutation-negative mild haemophilia A cases investigated showed normal gene dosage results by MLPA analysis and failed to identify the presence of any large deletions or duplications.

4.3.1 Severe Haemophilia A patients

In seven of the severe patients, deletions of 1 to 8 exons of *F8* were identified. MLPA confirmed deletions of exons 16-22, 7, 15-22, 24-25, 24, 7-10, and 7-9 in patients 5, 7, 9, 11, 12, 14 and 15 respectively. All of these deletions have been previously associated with a severe phenotype according to the HAMSTERS database, although not all of the patients in this study previously had the breakpoints of their mutations determined therefore it is not known whether some of them are exactly the same as those previously reported.

4.3.1.1 Application of direct assays

Prior to the introduction of MLPA, the mapping of breakpoints of large deletions allowed the development of direct screening assays to determine the carrier status of female relatives. Patient 5 is an example of an index case who had previously had the breakpoints of his large deletion mapped which allowed carriership to be determined in his sister (see Results section 3.3.2.1). Initial amplification of the essential regions of the gene had failed to amplify exons 16 to 21. A long PCR was then designed utilising a primer pair which amplified exons 15 through to 22. This produced a much smaller fragment than expected which indicated the presence of a partial deletion between introns 15-21. DNA sequencing was performed on the smaller fragment across the breakpoint which confirmed the identity of the product and indicated that the deletion removed exons 16 to 21. Long PCR encompassing exons 15 and 22 was then applied to his sister (Patient 6) which amplified the normal fragment and confirmed non-carriership of the large deletion. MLPA analysis confirmed both of the previous results in these patients.

In the case of Patient 7, it had not been previously possible to confirm the carrier status of his mother (Patient 8) using routine screening methods. Initial sequencing of *F8* had failed to amplify exon 7 in Patient 7. Repeat sequencing analysis indirectly confirmed this but as attempts to map this region failed a possible genotype of a suspected hemizygous deletion involving exon 7 was reported. Carrier diagnosis for the mother was therefore not possible as the precise nature of the mutation had not been confirmed and the deletion would have been masked by the normal allele on the other chromosome. MLPA analysis was able to confirm the exon 7 deletion originally identified by routine screening in Patient 7 and demonstrate its heterozygous presence in his carrier mother (see Results section 3.3.2.2)

This was also the case for Patient 9 who had a large deletion spanning exons 15-22. A long PCR was performed using primers for exons 14-23 which produced a smaller truncated fragment than expected. The boundaries of the deletion breakpoint were never mapped but the carrier diagnosis of his niece (Patient 10) was performed using the same primer pair. The truncated product was not detected which was suggestive of non-carriership but this needed to be corroborated by a direct screening assay. MLPA analysis confirmed the previous results for both patients and was an independent validation of MLPA results already obtained by the Sheffield Haemophilia service (see Results section 3.3.2.3).

4.3.1.2 Determination of carrier status

MLPA can be used to determine the carrier status in females when there is no index case available. An example of this is Patient 13 who was the mother of a severe haemophiliac. A deletion of exon 5 had been previously reported in her severe haemophiliac son. Specific analysis had been previously carried out on Patient 13 and she was shown to be potential carrier for the familial deletion but this required corroboration by a direct assay (see Results section 3.3.2.6). In this study, MLPA analysis was only performed on the mother as the sample for the son had been depleted. The heterozygous deletion of exon 5 was demonstrated in the mother, consistent with carriership. Disregarding the fact that the specific familial deletion had been previously determined, this

demonstrates that MLPA can reliably identify a large heterozygous deletion in a carrier in the absence of an index case. It can also be used to identify sporadic haemophilia carriers where 'de novo' mutations arise in the male germ cells. This was illustrated in a study by Lannoy, Abinet et al. (2009) where two girls with isolated FVIII deficiency and no history of familial haemophilia A were found to carry two different large gene deletions. They concluded that, in patients where evidence of inversion mutations involving intron 22 and 1 and point mutations have been eliminated, the presence of a *F8* deletion should be investigated by MLPA analysis.

4.3.1.3 Identification of large duplications

Until recently, large duplications had been rarely identified as the underlying cause of haemophilia A. This may have been due to the fact that whole-exon duplications were difficult to detect by PCR-based methods, but now with the development of MLPA, the presence of these rearrangements can be readily detected. In this study, MLPA data obtained for Patient 16 were suggestive of the presence of a large novel duplication of exon 1-22 combined with a single deletion of exon 25. Patient 16 (now deceased) was a severe female haemophiliac who may have been compound heterozygous for two haemophilia associated alleles which would correspond with the results obtained by MLPA. Her father was a haemophiliac and her mother was an obligate carrier of haemophilia A, however the mutation mechanisms in both had never been investigated. Interestingly, MLPA analysis for her severe haemophiliac nephew (Patient 17, also deceased) showed normal dosage quotients for all exons within *F8*, which confirmed the results previously obtained from the Newcastle haemophilia service (see Results section 3.3.2.9). Therefore the underlying genetic mechanism for the haemophilia phenotype within this patient was never identified. In this case, the MLPA results obtained for Patient 16 were undermined by the MLPA results in the index case (Patient 17). Although the samples for both patients passed quality control, they had been taken more than 10 years before so the sample quality would have to be questioned. The only way to resolve this case would be to repeat analysis on fresh samples, however as both patients were deceased this is not possible.

Samples obtained from living family members may possibly uncover the genetic mechanism within this family.

It is predicted that many duplications will result in a frame-shift of the *F8* mRNA and protein sequence if located in tandem orientation which is consistent with the severe phenotype of the disease (Rost, Loffler et al. 2008). In the study by Rost et al. (2008), they detected seven novel large duplications and one recently described duplication of exon 13 within *F8*. Large duplications have been reported for other genes and unequal crossing-over between Alu repeats had been described for the Duchenne/Becker muscular dystrophies (DMD/BMD) (Hu, Ray et al. 1991). The intronic sequences of *F8* are very rich in Alu repeats, especially introns 22 and 25. It has now been illustrated by Zimmerman et al. (2010), that repetitive elements such as LINEs and SINEs are usually located close to the breakpoint of *F8* duplications.

4.3.1.4 Presence of large deletions and link with development of inhibitors

Inhibitor development to the FVIII protein represents the major complication in patient care and can render classical substitution therapy ineffective. Inhibitor formation occurs at a frequency of 20-30% in severe haemophilia A, but is seen less frequently in mild and moderate cases. The evaluation of inhibitor risk and management of individuals is vital to lower morbidity and mortality rates of haemophilia A. Genetic factors such as particular types of mutations have been associated with a higher risk of developing an inhibitor e.g. intron 1 and 22 inversions and a small percentage of large deletions which are estimated to occur in ~50% of severe haemophilia A patients (Miller, Benson et al. 2012) .

The inhibitor status was determined for each index case in this study (apart from one deceased case which was unknown). All of the severe haemophilia A patients had developed an inhibitor whilst the mild cases had not. The prevalence of an inhibitor was 100% in those with a large deletion (7/7). Therefore results from this study corroborate the observation that the presence of a large deletion is a highly significant risk factor for the development of an inhibitor in severe haemophilia A (Oldenburg, El-Maarri et al. 2002). This is because mutations such as large deletions result in the absence or severe truncation of the FVIII protein which indicates that a key motivator in inhibitor

development is the presentation of a novel antigen to the patient's immune system (Oldenburg, Schroder et al. 2004).

The only patient found to have a possible large duplication was a severe female haemophiliac with no history of inhibitor development. It has been previously reported that patients with large duplications were also inhibitor-negative (Rost, Loffler et al. 2008). In this study, inhibitor risk in patients with no mutation was 0%, which is in agreement with a previous study which found a low inhibitor risk (~5%) in those without a mutation identifiable (Goodeve 2003).

4.3.2 Mild Haemophilia A cases

In all four mild haemophilia A cases investigated, FVIII:C levels had been reported as >5%, which occurs in 30-40% of cases of haemophilia A. The fact that MLPA analysis did not detect any large deletions or duplications was to be expected, as it has been shown that large deletions are not generally responsible for the mild form of the disease. A study by Bogdanova et al. (2007) investigated the spectrum of mutations and mutation detection rate in patients with mild and moderate haemophilia. Through sequence analysis they found that the vast majority of mutations within mild/moderate cases of haemophilia A are represented by missense mutations. In the mildly affected patients they accounted for ~76% of all cases, although other studies have shown that they can account for up to 86% (Schwaab, Oldenburg et al. 1995; Margaglione, Castaman et al. 2008; Castaman, Mancuso et al. 2009). These mutations result in the production of an abnormal FVIII protein or its reduced synthesis and/or release into the circulation (Jacquemin, De Maeyer et al. 2003). Although the mutations are spread throughout the entire sequence of *F8*, they are predominantly found in the A2 (mainly exon 11 and the A2/B boundary) and C1 (mainly exon 23) domains. Splicing errors were also found to be the cause of 9% of mild haemophilia A cases. The mutation detection rate in the mild cases investigated in this study was ~88% compared to mutation detection rate in severely affected males which was >95% in the study by Bogdanova, Markoff et al. (2005). Therefore analysis on the genomic level fails to detect the underlying molecular defect in up to 12% of mildly affected patients. Other studies have been published recently which report similar

mutation detection rates (Citron, Godmilow et al. 2002; Cutler, Mitchell et al. 2002; Liu, Nakaya et al. 2002; Habart, Kalabova et al. 2003).

Specific variants of coagulation proteins that directly interact with FVIII such as VWF, FII, FIX and FX may affect the amount and function of FVIII. In some cases mutations in *VWF* associated with the VWD variant type 2N, have an impact on the level of FVIII in the circulation and are responsible for the bleeding phenotype seen in these patients, indistinguishable from the one dependent upon mutations in *F8* itself. Both disorders display reduced levels of FVIII:C (~5-40 IU/dL) and normal-to-borderline-low levels of VWF. Three of the four mild haemophilia A patients investigated in this study had also previously been screened for type 2N associated mutations. All three were found to not to have any mutations associated with type 2N VWD therefore this theory could be excluded. However for the fourth patient (Patient 1) this could possibly be a second line of investigation.

In addition, variations in other proteins involved in folding, secretion, trafficking or stability of FVIII can lead to inherited bleeding disorders which could be mistaken for mild haemophilia A, such as F5F8D (see Introduction section 1.4.3.2). The combined F5F8D exhibits a mild to moderate bleeding tendency (Zhang and Ginsburg 2004). It is completely independent of FV deficiency and FVIII deficiency and arises because of mutations in two genes (*LMAN-1* and *MCFD2*). F5F8D is extremely rare but has an increased frequency in populations where there is a high incidence of consanguineous marriages. Only one of the four mild haemophilia A patients investigated in this study (Patient 4) had FV:C levels measured which may have been carried out in relation to his ethnic origin. Patient 4 had a FV:C level of 193 IU/dL which excluded the possibility of a F5F8D. His brother (Patient 3) had no recorded FV:C level and this may have been based on the test results of patient 4. However this is another area for investigation to consider for patients with no detectable mutation.

The fact that haemophilia A patients exist with no identifiable *F8*, *VWF*, or known loci involved in combined defects such as F5F8D indicates that there

are other unidentified loci that may be involved in the disease phenotype. The four mutation-negative patients were all male, suggesting X-chromosomal inheritance. Therefore the underlying genetic defect may be associated with the *F8* locus itself or in some kind of localised interaction with *F8* or the FVIII protein. The location of these mutations could possibly be in regions which are not routinely scanned or that may be undetectable by current methods, in the patients in this study this encompasses both DNA sequencing of the essential regions of *F8* and MLPA analysis. One theory is that some patients with no mutation identifiable in the coding sequence and proximal intronic regions could be affected by either distant intronic variations impacting on splicing efficiency or duplications of certain exons which may be detectable at the RNA level. A mRNA-based approach has been adopted in studies attempting to uncover the mutations in severe haemophilia A patients with no detectable mutation (El-Maarri, Herbiniaux et al. 2005; El-Maarri, Singer et al. 2006). Apart from the discovery of rare mutations resulting in a lack of mRNA production, these attempts have shown disappointing results and this method has rarely been implemented in patients with mild haemophilia A (Castaman, Giacomelli et al. 2010). A recent study by Castaman and Toso (2011) investigated a number of mild haemophilia A patients in whom no *F8* mutations had been identified and who also had a reduced biologic response to the infusion of desmopressin. In five of the six patients investigated, three different intronic mutations were identified after sequencing the introns associated with the observed *F8* mRNA abnormality. The mild phenotype was explained by the fact that mRNA evaluation indicated the presence of a normal mRNA transcript as well as the aberrant mRNA. These variations were not identified in normal controls and were always present as heterozygous traits in the carrier female relatives of the patients. The study showed that alternative pathologic mRNA splicing in mild haemophilia A may be caused by deep intronic point variations which are probably responsible for abnormal mRNA processing forming the basis of disease in these patients. Therefore this should be considered for the small number of patients in whom no mutation can be identified although these aberrations need to be further characterised before they can be used in genetic diagnosis.

In these cases, MLPA would not be expected to detect the mutation as these different mutation mechanisms are not suited to the MLPA approach that has been adopted in this current study.

4.4 VWD PATIENTS

In total, twenty-six VWD patients were investigated by MLPA analysis in this study. Ten of the twenty-six patients had been diagnosed with type 3 VWD and had undergone genetic investigation as part of a type 3 VWD study at the MDC. Seven of the ten type 3 VWD patients had been previously found to have the novel exon 4-5 deletion (c.221-977_532+7059del mutation) discovered in the study. Three of the seven were also previously found to be compound heterozygous for the exon 4-5 deletion and another mutation. MLPA confirmed the presence of the large deletion in all seven patients. The remaining three patients without the large deletion had been previously found to have a small deletion, duplication, or point mutation. MLPA analysis eliminated the presence of a large deletion in these three cases, however a small point mutation previously found in one patient (T3-9) illustrated how this caused dropout of the MLPA probe signal for one exon (see section 4.2.5). Patient T3-9 had been previously found to show compound heterozygosity for two DNA single base changes (c.5200C>T and c.7730-3C>G). Although neither of these *VWF* mutations have been previously reported, due to their nature they are highly likely to account for the severe VWD phenotype seen in this patient. Therefore this patient would not be considered suitable for MLPA analysis, however they were included in the study to illustrate how a mutation underneath the MLPA probe can affect the probe signal.

The remaining sixteen patients had been diagnosed with type 1 VWD and had been previously recruited into a UK national study of type 1 VWD for genetic investigation. Two of the sixteen type 1 VWD patients had previously been identified to be heterozygous for the exon 4-5 deletion as part of the type 3 VWD study. This was confirmed by MLPA analysis in this study. The remaining fourteen type 1 VWD patients had previously been found to have no causative mutation responsible for their phenotype. No large deletions were identified by MLPA analysis in this group of patients.

4.4.1 Confirmation of the Exon 4-5 deletion in VWD patients

Large deletions are considered to be a rare cause of VWD, which generally result in a complete absence of VWF protein (Schneppenheim, Castaman et al. 2007). So far, only twelve large deletions have been reported in association with type 3 VWD, varying in size from a single exon to the entire gene. An additional large deletion has been described for type 2A, which spans all VWF A domains (www.vwf.group.shef.ac.uk). The exon 4-5 deletion was first discovered by Sutherland et al. (2009) in a group of patients from the North West of England using direct sequencing of *VWF* at the genomic DNA level and of *VWF* mRNA derived cDNA. In this study, MLPA analysis confirmed the presence of the exon 4-5 deletion in the same cohort of patients originally investigated in the type 3 study.

4.4.2 Confirmation of heterozygosity for the exon 4-5 deletion

Large deletions in *VWF* are easily identified in homozygous VWD patients by an absence of PCR products of the deleted fragment. However conventional PCR methods cannot identify heterozygous large deletions in relatives due to the amplification of exons present in the normal chromosome. Of the seven type 3 VWD patients found to have the deletion, three were found to be homozygous for the mutation, and the remaining four patients were heterozygous for the deletion. Three of the four patients heterozygous for the deletion were also compound heterozygous for a second *VWF* mutation, which explained their type 3 VWD phenotype. MLPA was clearly able to confirm the homozygous or heterozygous nature of the deletion in all of these individuals but did not identify compound heterozygosity for the second *VWF* mutation seen in some patients. This validates one of the limitations of MLPA in that it is not suitable for identifying unknown point mutations. Therefore in the case of compound heterozygosity for a large deletion and a point mutation, both direct sequencing and MLPA analysis would be required in order to determine the underlying genetic defect in a patient.

4.4.3 Confirmation of the exon 4-5 deletion in type 1 VWD patients

Sutherland et al. (2009) had investigated the possibility of the occurrence of the exon 4-5 deletion in type 1 VWD. This was due to the high overall frequency of

the deletion mutation in the type 3 VWD cohort, and the fact that heterozygosity for this mutation may not be identified by conventional analysis at the genomic DNA level. The two index type 1 cases (T1-3 and T1-23) identified as being heterozygous for the exon 4-5 deletion were chosen for investigation by MLPA in this study. Of the two cases one was heterozygous for the mutation (T1-23) whilst the other was compound heterozygous for a second *VWF* mutation (T1-3). MLPA confirmed the heterozygous nature of the deletion in both individuals but did not identify compound heterozygosity for the second *VWF* mutation seen in patient T1-3.

The fact that the exon 4-5 deletion was identified in both type 1 and type 3 VWD patients is unusual as both types have different inheritance patterns and genetics. Type 3 is inherited in an autosomal recessive manner and *VWF* mutations are usually present in the homozygous or compound heterozygous state in individuals with the disease phenotype. Individuals with heterozygous mutations are usually asymptomatic. However it has been reported that when certain mutations associated with type 3 VWD are inherited in the heterozygous state, they have the ability to reduce plasma VWF levels to those consistent with a type 1 VWD phenotype (Anvret, Blomback et al. 1992; Schneppenheim, Krey et al. 1994; Zhang, Blomback et al. 1994; Lethagen, Isaksson et al. 2002). Expression studies (performed by Mackenzie Bowman, Queen's University, Kingston, Canada) indicate that heterozygosity for the exon 4-5 deletion leads to defective multimerisation and decreased secretion of the mutant VWF protein produced, resulting in a marked reduction in normal VWF multimers present in the circulation. This would be consistent with the phenotype seen in the two type 1 VWD index cases studied and supports the theory of a dominant-negative effect of the deletion and its association with type 1 VWD (Sutherland et al. 2009).

4.4.4 Overview of the exon 4-5 deletion

All of the individuals found to have the exon 4-5 deletion were of British Caucasian origin. Haplotype analysis by Sutherland et al. (2009) demonstrated that in the majority of cases the deletion originated from a single founder event which originated in Britain. However there was evidence for at least one

independent occurrence or a possible recombination event. Therefore combined with the proposed mechanism for the deletion (see Introduction section 1.14.3), it is possible that the exon 4-5 deletion may not be confined to British Caucasians (Sutherland et al. 2009). A recent study by Bowman et al. (2013) investigated the mutational spectrum of a cohort of Canadian type 3 VWD patients. They identified one index case as being homozygous for the exon 4-5 deletion, which they confirmed using the deletion-specific PCR described by Sutherland et al. (2009) and MLPA. They also found that two family members enrolled in the study, a sibling and mother, diagnosed with type 1 VWD were heterozygous for the exon 4-5 deletion. Analysis of the haplotype of these individuals was found to be different than the common haplotype reported in the UK population (Sutherland et al. 2009). Likewise, in the Canadian and European type 1 VWD studies, no common haplotype was found between individuals with the same mutations (Goodeve 2007; James, Notley et al. 2007). Another founder mutation which has been identified in type 3 VWD is the single cytosine deletion in exon 18 of *VWF* (c.2435delC) which has been shown to be prevalent in the Swedish and Dutch populations due to its mechanism (Eikenboom, Ploos van Amstel et al. 1992; Zhang, Falk et al. 1992; Zhang, Blomback et al. 1993; Zhang, Blomback et al. 1994).

MLPA is ideally suited to the detection of large deletions such as the exon 4-5 deletion, once point mutations have been eliminated, however mutations such as the single cytosine deletion will be unsuitable for detection by MLPA. The current probe kit for *VWF* does not distinguish it, therefore a specific MLPA probe would have to be developed in order to identify the mutation.

4.4.5 Type 1 VWD with no causative mutation

MLPA analysis of the fourteen remaining type 1 VWD patients failed to identify any large deletions within *VWF* that may have been responsible for the molecular pathogenesis of their disease. This was to be expected as large deletions are considered to be a rare cause of VWD and are more usually associated with an autosomal recessive mode of inheritance, meaning that individuals heterozygous for such deletions are likely to be asymptomatic. This

meant that overall no causative *VWF* mutation had been found to account for the type 1 VWD phenotype present in these patients.

Previous DNA sequencing studies had identified *VWF* polymorphisms in nine of the fourteen type 1 VWD patients. The highly polymorphic nature of *VWF* could potentially lead to defective binding of one or more MLPA probes to their target sequences, causing a reduction in probe peak height which could be misinterpreted as a deletion (see section 4.2.5). The *VWF* polymorphism p.Tyr1584Cys, a missense mutation c.4751A>G in exon 28, had been previously identified in four of the nine patients. This was not found to have an effect on the peak height of the probe for exon 28 and therefore its presence was not evident by MLPA. In the remaining five patients it was difficult to ascertain whether some of the putative *VWF* polymorphisms identified by DNA sequencing in some of the patients had an effect on probe height. This was due to the fact the majority of the exons in which they were present were part of the *VWF* kit 2 which consistently failed for these patients.

In type 1 VWD, approximately 75% of mutations are missense, some of which may result in single amino acid substitutions with a dominant negative effect that may interfere with the intracellular transport of *VWF* subunits, leading to a reduction in *VWF* secretion ((Haberichter, Castaman et al. 2008). Recent studies have identified other mutations that result in the rapid clearance of *VWF* from the circulation. In the case of these mutations, *VWF* is synthesised and secreted normally but is removed from the circulation more quickly than expected (Sztukowska, Gallinaro et al. 2008; Sadler 2009). The impaired secretion and increased clearance are most likely the common molecular mechanisms that lead to type 1 VWD.

As indicated above, the majority of point mutations would not be routinely detected by MLPA and the application of the method should only be employed after elimination of point mutations by DNA sequencing. Exceptions may include the application of MLPA to the detection of known deletions prevalent in local populations, such as the exon 4-5 deletion discussed.

Type 1 VWD accounts for more than 70% of VWD cases and is regarded as the most frequent form of VWD. However due to a number of factors, it is not unusual that in many cases a causative *VWF* mutation is not found to account for the type 1 VWD diagnosis. A number of environmental factors such as age, stress and pregnancy can lead to variations in *VWF* levels as well as genetic modifiers at loci away from *VWF*, such as ABO blood group (Ginsburg and Bowie 1992; Ginsburg 1999; Keightley, Lam et al. 1999; Levy and Ginsburg 2001; Laffan, Brown et al. 2004). In this study, eleven of the fourteen patients with no causative mutation identified were of blood group O (see Results, Table 3.5). If individuals are blood group O with the polymorphism p.Tyr1584Cys as in the case of patients T1-16, T1-19 and T1-25, this may lead to decreased *VWF* levels and increased susceptibility to bleeding (Bowen 2003; O'Brien, James et al. 2003; Bowen, Collins et al. 2005; Bowen and Collins 2006). The other polymorphisms previously identified in the type 1 VWD patients (see Results, Table 3.5) including p.Arg924Gln, c.5834-8C>G, p.Gly2705Arg, p.Val1850Met, c.5170+10C>T, p.Asp141Gly and p.Thr2647Met are not thought likely to directly cause the disease phenotype in these patients as they did not segregate with the disease in the families in the initial study (Cumming, Grundy et al. 2006; Hickson, Hampshire et al. 2010). Although, it has been reported that the presence of the polymorphism p.Arg924Gln will result in a reduction in *VWF* levels, especially if in association with blood group O, however single inheritance of this polymorphism may be insufficient for VWD diagnosis (Hickson, Hampshire et al. 2010).

Therefore in conclusion, type 1 VWD is a disorder that is likely to be influenced by genetic modifiers for which the current MLPA approach is not suited. MLPA analysis has not made a significant contribution to identifying the underlying genetic defect within this group of type 1 VWD patients, and further investigations would need to be carried out to reveal the pathogenic mechanisms associated with their disease. Genome wide linkage could possibly be used to identify genetic loci away from *VWF* itself that influence *VWF:Ag* levels.

4.5 MUTATION DETECTION STRATEGIES – THE ROLE OF MLPA

The molecular diagnosis of both haemophilia A and VWD follow practice guidelines that have been written by the UKHCDO. MLPA is a technique which has been recommended to be used in addition to gold standard techniques already employed to overcome any limitations they may have. This study was carried out in order to assess how MLPA will fit into the currently used mutation screening protocols for both haemophilia A and VWD within the MDC.

4.5.1 Proposed Mutation Screening Algorithm incorporating MLPA for Haemophilia A and VWD

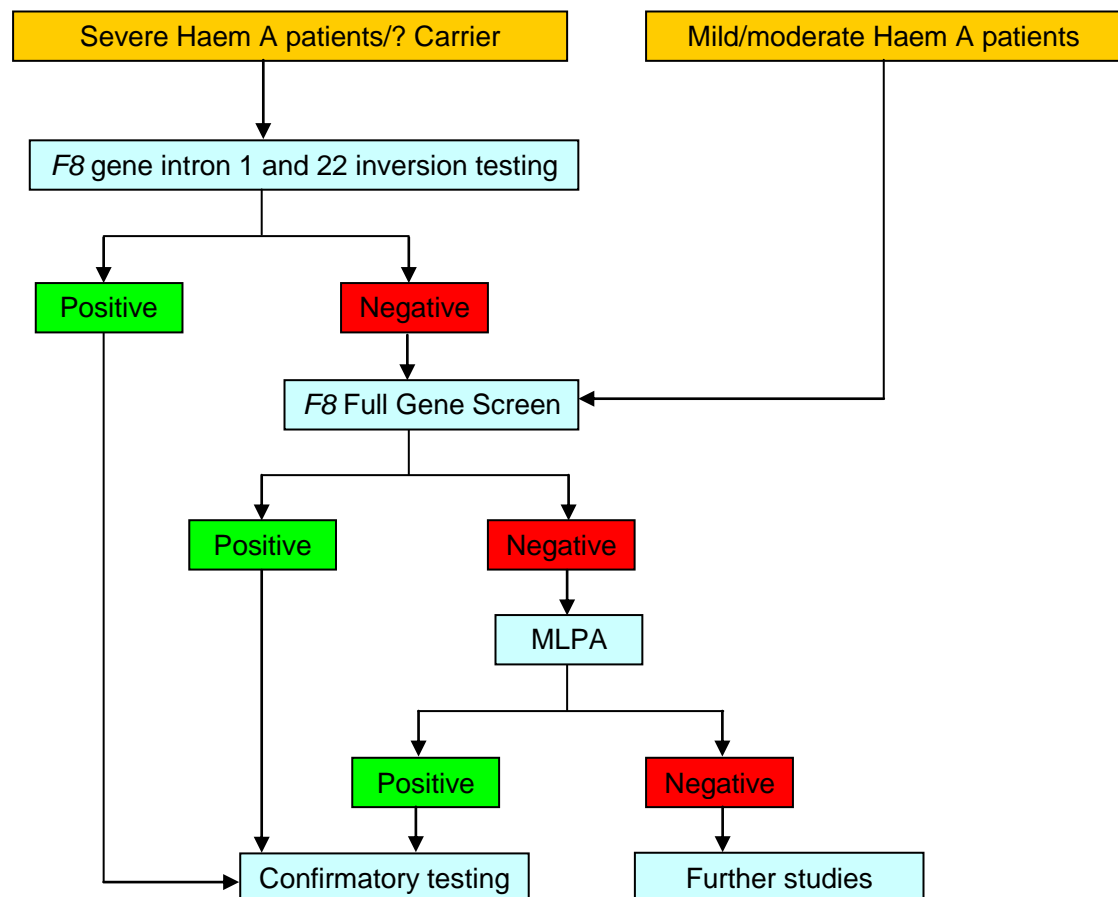
4.5.1.1 Haemophilia A

The severity of the haemophilia A in the family determines the diagnostic strategy which is applied. Severe haemophiliacs are first screened for the *F8* intron 22 inversion mutation followed by the *F8* intron 1 inversion mutation which should identify the underlying mutation in 45-50% of patients. The remaining severe haemophilia A families should then have full mutation analysis of *F8* carried out. As a common mutational mechanism is not associated with mild or moderate haemophilia A, full mutation analysis is required in all cases.

The gold standard for mutation detection in affected males, or in obligate carrier females where an affected male is not available, is DNA sequencing. Although DNA sequencing has a very high level of sensitivity, a small proportion of patients have a mutation which lies outside of the regions being analysed which will go undetected. The current DNA sequencing strategies have been shown to detect mutations or candidate mutations in 98% of severe haemophilia A males (Klopp, Oldenburg et al. 2002). However, approximately 2-18% of patients overall have not been found to have a causative mutation (Oldenburg, Ivaskevicius et al. 2001; Jayandharan, Shaji et al. 2005; Bogdanova, Markoff et al. 2007). Some of these cases may have mutations which are located deep within introns and therefore may only be detectable at the RNA level (Bagnall, Waseem et al. 1999) or be mediated by mutation mechanisms yet to be explained (El-Maarri, Herbiniaux et al. 2005). DNA sequencing may also miss mutation events such as sequence duplications (Acquila, Pasino et al. 2004;

Rost, Loffler et al. 2008) and heterozygous large deletions where a female carrier is used for initial mutation screening. Therefore if no mutation is identified through sequencing, MLPA should be performed in order to detect the presence of possible large deletions/duplications responsible for the disease (see Figure 4.1). Large deletions are associated with an increased risk for the development of inhibitors, therefore the introduction of MLPA will also help to classify a higher proportion of those at high risk, should it identify deletion mutations that were masked previously. Following positive identification of a large deletion/duplication, results from the initial assay should always be confirmed by repeat testing with an independent method where this is possible. For those with no mutation identified, further investigations such as mRNA analysis may be carried out.

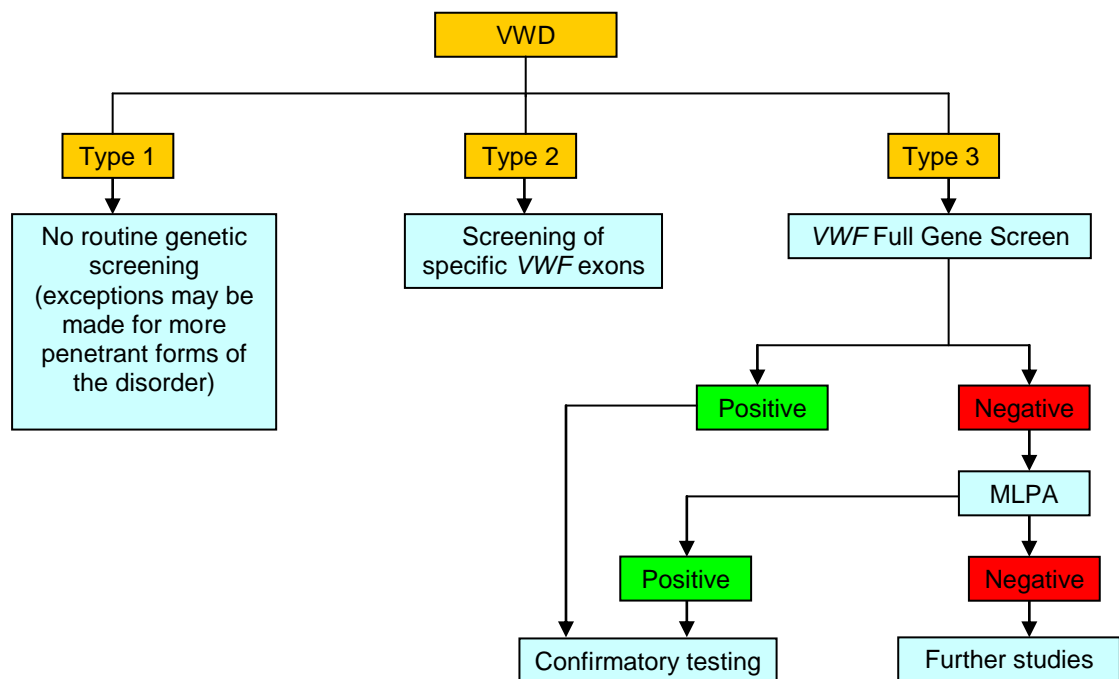
Figure 4.1: Mutation screening algorithm incorporating MLPA for haemophilia A patients



4.5.1.2 VWD

The classification of VWD is initially based on clinical and phenotypic data, which is fundamental to the investigation of the disorder. Genetic testing is not always required in VWD and varies for the different VWD subtypes e.g. Type 2 requires screening of specific functional domains represented by exons 18-20 and exon 28. In type 3 VWD, genetic testing is beneficial due to the severe bleeding phenotype and autosomal recessive pattern of inheritance. No universal mutation hotspots have been identified in type 3 VWD, therefore mutation analysis is carried out by direct DNA sequencing of the essential regions of *VWF*. Mutation pre-screening methods are not useful in VWD analysis due to the highly polymorphic nature of *VWF* which can result in unnecessary sequencing of amplicons (Hashemi et al. 2007). If no mutations are identified through sequencing, or if a patient appears to be homozygous for a given mutation, the prospect of an insertion, rearrangement or whole or partial gene deletion on the other *VWF* allele should be taken into account. As in the case of haemophilia A, MLPA should then be performed in order to detect the presence of possible large deletions, with positive results confirmed by an independent method, where possible. Further studies may be required for mutation-negative cases (see Figure 4.2).

Figure 4.2: Mutation screening algorithm incorporating MLPA for VWD patients



4.6 CONCLUSIONS AND FUTURE WORK

The results of this study show that MLPA, once it has been optimised within the laboratory, can be applied to the investigation of large deletions/duplications in haemophilia A and VWD patients and that the method can be easily incorporated into a diagnostic genetic screening service. MLPA provides a streamlined approach to the identification of genetic variants which previously were either difficult to detect or were masked and allows for a more comprehensive explanation of the molecular pathogenesis of the disease in a greater number of index cases. However some cases still remain unexplained following extensive mutation analysis which suggests that there may be undetermined mechanisms underlying a small number of cases or possible misdiagnosis.

In haemophilia A, for the small number of cases in whom no mutation can be identified, the bleeding phenotype may be explained by mutations in *VWF* that affects the level of FVIII in these patients (type 2N VWD), which produces a disease phenotype which is very similar to the one reliant on mutations in *F8*

itself. Also mutations within the *LMAN1* and *MCFD2* genes result in combined FVIII and FV deficiency and produce a FVIII deficiency phenotype (Zhang, McGee et al. 2006). However in haemophilia A cases where no relevant mutations can be identified in *F8*, or implicated loci including *VWF*, *LMAN1* and *MCFD2*, suggests that the underlying genetic mutations or modifiers are either located in regions that are yet to be identified or are of a type that has not been identified. It has been shown that deep intronic point variations which are probably responsible for abnormal mRNA processing may form the basis of disease in a proportion of these patients. Therefore future work should consider mRNA studies in these patients or potentially using next generation sequencing techniques to sequence all intronic and exonic regions for mutations that might lie deep within the introns or to identify further loci which may modify FVIII levels in these individuals. This would require further characterisation before being considered for genetic diagnosis.

MLPA kits are also available for the factor 9 gene (*F9*) and therefore would be useful for identifying deletions/duplications responsible for haemophilia B. The number of haemophilia B investigations performed by the MDC is lower and to date no patients investigated for the presence of *F9* mutations have had unexplained phenotype at the genetic level. If there was a requirement for this within the genetic service at the MDC it would be very easy to set up this assay as the only difference would be the probemix and therefore would only require validation of results.

Carrier and prenatal diagnosis are often requested in families of haemophilia A and type 3 VWD. Once a large deletion/duplication has been characterised in an individual by MLPA, a direct assay can be designed in most cases, if the breakpoints of the mutation have been established. Further work is required for some haemophilia A patients in this study to characterise their deletions by breakpoint mapping and develop specific assays for use within their families. If the mutation has not been characterised in a particular family, by all methods available then segregation analysis using polymorphic markers may be considered (Hallden, Knobe et al. 2012). However, these assume linkage to the gene and cannot be employed in sporadic cases.

In type 3 VWD, a causative *VWF* mutation(s) can be identified in the majority of cases to account for the bleeding phenotype, however there are cases when the phenotype cannot be fully explained by the genotype using conventional DNA screening methods. It has been shown by Sutherland et al. (2009) that RNA analysis methods may be used to identify aberrant splicing events or large heterozygous deletions which are masked during DNA analysis. The novel exon 4-5 deletion mutation was identified and found to cause type 3 VWD, when inherited in homozygosity or compound heterozygosity, and type 1 VWD when inherited in heterozygosity (Sutherland et al. 2009). The MLPA technique was successfully employed to detect this large deletion in both VWD subtypes. However, there are cases of type 3 VWD where the genetic cause remains unclear. This could point to unexplained underlying mechanisms such as apparently silent sequence variations in *VWF* located outside of consensus splice sites that disrupt the normal *VWF* mRNA splicing, deep intronic mutations or distant regulatory elements outside of *VWF*. In some cases it may even suggest a revision of diagnoses based on historical phenotypic data which may be inaccurate. Future work involving expression studies or whole gene sequencing in these individuals may enable the fully comprehensive understanding of the molecular basis of type 3 VWD.

The causative molecular defect in type 1 VWD underlying the disorder is unknown in a significant number of cases and even in those cases where a causative *VWF* mutation is known, the associated molecular pathology is not necessarily understood. MLPA analysis did not detect the presence of heterozygous large deletions in this cohort of patients (apart from the exon 4-5 deletion previously identified in 2 index cases), therefore failing to further elucidate the underlying genetic defect in these individuals. This supports the principle that there is limited clinical use for genetic diagnosis in type 1 VWD (Keeney and Cumming 2001; Keeney, Bowen et al. 2008). Future work to further understand the molecular basis of type 1 VWD may involve genome wide linkage analysis to search for modifying genes or factors away from the *VWF* locus that affect *VWF* levels.

This project has evaluated the application of MLPA in a diagnostic setting for the bleeding disorder service covering the North West (NW) of England. This has shown that it can be considered to be a useful diagnostic tool when conventional mutation screening has failed to identify a causative mutation, as long as its limitations are understood and where possible it is not relied upon solely.

REFERENCES

- ACQUILA, M., PASINO, M., LANZA, T., BOTTINI, F., MOLINARI, A. C. & BIOCCHI, M. P. 2004. Duplication of exon 13 causes one third of the cases of mild hemophilia A in northern Italy. *Haematologica*, 89, 758-9.
- ALLEN, S., ABUZENADAH, A. M., HINKS, J., BLAGG, J. L., GURSEL, T., INGERSLEV, J., GOODEVE, A. C., PEAKE, I. R. & DALY, M. E. 2000. A novel von Willebrand disease-causing mutation (Arg273Trp) in the von Willebrand factor propeptide that results in defective multimerization and secretion. *Blood*, 96, 560-8.
- ANTONARAKIS, S. E. 1998. Molecular genetics of coagulation factor VIII gene and haemophilia A. *Haemophilia*, 4 Suppl 2, 1-11.
- ANVRET, M., BLOMBACK, M., LINDSTEDT, M., SODERLIND, E., TAPPER-PERSSON, M. & THELANDER, A. C. 1992. Genetic and blood coagulation characterization of "Swedish" families with von Willebrand's disease types I and III: new aspects of heredity. *Hum Genet*, 89, 147-54.
- BAGNALL, R. D., WASEEM, N. H., GREEN, P. M., COLVIN, B., LEE, C. & GIANNELLI, F. 1999. Creation of a novel donor splice site in intron 1 of the factor VIII gene leads to activation of a 191 bp cryptic exon in two haemophilia A patients. *Br J Haematol*, 107, 766-71.
- BAGNALL, R. D., WASEEM, N., GREEN, P. M. & GIANNELLI, F. 2002. Recurrent inversion breaking intron 1 of the factor VIII gene is a frequent cause of severe hemophilia A. *Blood*, 99, 168-74.
- BAHNAK, B. R., WU, Q. Y., COULOMBEL, L., ASSOULINE, Z., KERBIRIOU-NABIAS, D., PIETU, G., DROUET, L., CAEN, J. P. & MEYER, D. 1989. Expression of von Willebrand factor in porcine vessels: heterogeneity at the level of von Willebrand factor mRNA. *J Cell Physiol*, 138, 305-10.
- BARONCIANI, L., COZZI, G., CANCELANI, M. T., PEYVANDI, F., SRIVASTAVA, A., FEDERICI, A. B. & MANNUCCI, P. M. 2000. Molecular characterization of a multiethnic group of 21 patients with type 3 von Willebrand disease. *Thromb Haemost*, 84, 536-40.
- BARONCIANI, L., COZZI, G., CANCELANI, M. T., PEYVANDI, F., SRIVASTAVA, A., FEDERICI, A. B. & MANNUCCI, P. M. 2003. Molecular defects in type 3 von Willebrand disease: updated results from 40 multiethnic patients. *Blood Cells Mol Dis*, 30, 264-70.
- BARONCIANI, L., FEDERICI, A. B., COZZI, G., LA MARCA, S., PUNZO, M., RUBINI, V., CANCELANI, M. T. & MANNUCCI, P. M. 2008. Expression studies of missense mutations p.D141Y, p.C275S located in the propeptide of von Willebrand factor in patients with type 3 von Willebrand disease. *Haemophilia*, 14, 549-55.

BECKER, J., SCHWAAB, R., MOLLER-TAUBE, A., SCHWAAB, U., SCHMIDT, W., BRACKMANN, H. H., GRIMM, T., OLEK, K. & OLDENBURG, J. 1996. Characterization of the factor VIII defect in 147 patients with sporadic hemophilia A: family studies indicate a mutation type-dependent sex ratio of mutation frequencies. *Am J Hum Genet*, 58, 657-70.

BERLINER, S. A., SELIGSOHN, U., ZIVELIN, A., ZWANG, E. & SOFFERMAN, G. 1986. A relatively high frequency of severe (type III) von Willebrand's disease in Israel. *Br J Haematol*, 62, 535-43.

BERNARDI, F., MARCHETTI, G., GUERRA, S., CASONATO, A., GEMMATI, D., PATRACCHINI, P., BALLERINI, G. & CONCONI, F. 1990. A de novo and heterozygous gene deletion causing a variant of von Willebrand disease. *Blood*, 75, 677-83.

BOGDANOVA, N., MARKOFF, A., POLLMANN, H., NOWAK-GOTTL, U., EISERT, R., WERMES, C., TODOROVA, A., EIGEL, A., DWORNICZAK, B. & HORST, J. 2005. Spectrum of molecular defects and mutation detection rate in patients with severe hemophilia A. *Hum Mutat*, 26, 249-54.

BOGDANOVA, N., MARKOFF, A., EISERT, R., WERMES, C., POLLMANN, H., TODOROVA, A., CHLYSTUN, M., NOWAK-GOTTL, U. & HORST, J. 2007. Spectrum of molecular defects and mutation detection rate in patients with mild and moderate hemophilia A. *Hum Mutat*, 28, 54-60.

BOLTON-MAGGS, P. H. & PASI, K. J. 2003. Haemophilias A and B. *Lancet*, 361, 1801-9.

BONTHRON, D., ORR, E. C., MITSOCK, L. M., GINSBURG, D., HANDIN, R. I. & ORKIN, S. H. 1986. Nucleotide sequence of pre-pro-von Willebrand factor cDNA. *Nucleic Acids Res*, 14, 7125-7.

BOWEN, D. J. 2002. Haemophilia A and haemophilia B: molecular insights. *Mol Pathol*, 55, 127-44.

BOWEN, D. J. 2003. An influence of ABO blood group on the rate of proteolysis of von Willebrand factor by ADAMTS13. *J Thromb Haemost*, 1, 33-40.

BOWEN, D. J., COLLINS, P. W., LESTER, W., CUMMING, A. M., KEENEY, S., GRUNDY, P., ENAYAT, S. M., BOLTON-MAGGS, P. H., KEELING, D. M., KHAIR, K., TAIT, R. C., WILDE, J. T., PASI, K. J. & HILL, F. G. 2005. The prevalence of the cysteine1584 variant of von Willebrand factor is increased in type 1 von Willebrand disease: co-segregation with increased susceptibility to ADAMTS13 proteolysis but not clinical phenotype. *Br J Haematol*, 128, 830-6.

BOWEN, D. J. & COLLINS, P. W. 2006. Insights into von Willebrand factor proteolysis: clinical implications. *Br J Haematol*, 133, 457-67.

BOWMAN, M., TUTTLE, A., NOTLEY, C., BROWN, C., TINLIN, S., DEFOREST, M., LEGGO, J., BLANCHETTE, V.S., LILICRAP, D., JAMES, P. (2013). "The Genetics of Canadian Type 3 von Willebrand Disease (VWD): Further Evidence for Co-dominant Inheritance of Mutant Alleles." *J Thromb Haemost* **11**(3): 512-20.

BUDDE, U. & SCHNEPPENHEIM, R. 2001. von Willebrand factor and von Willebrand disease. *Rev Clin Exp Hematol*, 5, 335-68; quiz following 431.

CABRERA, N., CASANA, P., CID, A. R., MORET, A., MORENO, M., PALOMO, A. & AZNAR, J. A. 2011. First application of MLPA method in severe von Willebrand disease. Confirmation of a new large VWF gene deletion and identification of heterozygous carriers. *Br J Haematol*, 152, 240-2.

CASANA, P., MARTINEZ, F., ESPINOS, C., HAYA, S., LORENZO, J. I. & AZNAR, J. A. 1998. Search for mutations in a segment of the exon 28 of the human von Willebrand factor gene: new mutations, R1315C and R1341W, associated with type 2M and 2B variants. *Am J Hematol*, 59, 57-63.

CASANA, P., MARTINEZ, F., HAYA, S., LORENZO, J. I., ESPINOS, C. & AZNAR, J. A. 2000. Q1311X: a novel nonsense mutation of putative ancient origin in the von Willebrand factor gene. *Br J Haematol*, 111, 552-5.

CASONATO, A., GAUCHER, C., PONTARA, E., ZUCCHETTO, A., ZERBINATI, P., MAZURIER, C. & GIROLAMI, A. 1998. Type 2N von Willebrand disease due to Arg91Gln substitution and a cytosine deletion in exon 18 of the von Willebrand factor gene. *Br J Haematol*, 103, 39-41.

CASONATO, A., PONTARA, E., SARTORELLO, F., CATTINI, M. G., GALLINARO, L., BERTOMORO, A., ROSATO, A., PADRINI, R. & PAGNAN, A. 2006. Identifying type Vicenza von Willebrand disease. *J Lab Clin Med*, 147, 96-102.

CASONATO, A., PONTARA, E., SARTORELLO, F., CATTINI, M. G., PERUTELLI, P., BERTOMORO, A., GALLINARO, L. & PAGNAN, A. 2007. Identifying carriers of type 2N von Willebrand disease: procedures and significance. *Clin Appl Thromb Hemost*, 13, 194-200.

CASONATO, A., SARTORELLO, F., CATTINI, M. G., PONTARA, E., SOLDERA, C., BERTOMORO, A. & GIROLAMI, A. 2003. An Arg760Cys mutation in the consensus sequence of the von Willebrand factor propeptide cleavage site is responsible for a new von Willebrand disease variant. *Blood*, 101, 151-6.

CASTAMAN, G., EIKENBOOM, J. C., MISSIAGLIA, E. & RODEGHIRO, F. 2000. Autosomal dominant type 1 von willebrand disease due to G3639T mutation (C1130F) in exon 26 of von Willebrand factor gene: description of five Italian families and evidence for a founder effect. *Br J Haematol*, 108, 876-9.

CASTAMAN, G., GIACOMELLI, S. H., MANCUSO, M. E., SANNA, S., SANTAGOSTINO, E. & RODEGHIERO, F. 2010. F8 mRNA studies in haemophilia A patients with different splice site mutations. *Haemophilia*, 16, 786-90.

CASTAMAN, G., MANCUSO, M. E., GIACOMELLI, S. H., TOSETTO, A., SANTAGOSTINO, E., MANNUCCI, P. M. & RODEGHIERO, F. 2009. Molecular and phenotypic determinants of the response to desmopressin in adult patients with mild hemophilia A. *J Thromb Haemost*, 7, 1824-31.

CASTAMAN, G. & TOSETTO, A. 2011. Is a reduced intensity treatment with Rituximab effective in acquired haemophilia A? *Haemophilia*, 17, 817-8.

CATTANEO, M., FEDERICI, A. B., LECCHI, A., AGATI, B., LOMBARDI, R., STABILE, F. & BUCCIARELLI, P. 1999. Evaluation of the PFA-100 system in the diagnosis and therapeutic monitoring of patients with von Willebrand disease. *Thromb Haemost*, 82, 35-9.

CITRON, M., GODMILOW, L., GANGULY, T. & GANGULY, A. 2002. High throughput mutation screening of the factor VIII gene (F8C) in hemophilia A: 37 novel mutations and genotype-phenotype correlation. *Hum Mutat*, 20, 267-74.

COLLINS, P. W., CUMMING, A. M., GOODEVE, A. C. & LILLICRAP, D. 2008. Type 1 von Willebrand disease: application of emerging data to clinical practice. *Haemophilia*, 14, 685-96.

COONEY, K. A., NICHOLS, W. C., BRUCK, M. E., BAHOU, W. F., SHAPIRO, A. D., BOWIE, E. J., GRALNICK, H. R. & GINSBURG, D. 1991. The molecular defect in type IIB von Willebrand disease. Identification of four potential missense mutations within the putative Gplb binding domain. *J Clin Invest*, 87, 1227-33.

CUMMING, A., GRUNDY, P., KEENEY, S., LESTER, W., ENAYAT, S., GUILLIATT, A., BOWEN, D., PASI, J., KEELING, D., HILL, F., BOLTON-MAGGS, P. H., HAY, C. & COLLINS, P. 2006. An investigation of the von Willebrand factor genotype in UK patients diagnosed to have type 1 von Willebrand disease. *Thromb Haemost*, 96, 630-41.

CUMMING, A.M., SUTHERLAND, M.S., KEENEY, S. (2010) "The molecular basis of type 2 von Willebrand disease." *Journal of Coagulation Disorders*, 2, 29-35.

CUMMING, A. M. 2004. The factor VIII gene intron 1 inversion mutation: prevalence in severe hemophilia A patients in the UK. *J Thromb Haemost*, 2, 205-6.

CUMMING, A. M. & WENSLEY, R. T. 1993. Analysis of von Willebrand factor multimers using a commercially available enhanced chemiluminescence kit. *J Clin Pathol*, 46, 470-3.

CUTLER, J. A., MITCHELL, M. J., SMITH, M. P. & SAVIDGE, G. F. 2002. The identification and classification of 41 novel mutations in the factor VIII gene (F8C). *Hum Mutat*, 19, 274-8.

DAVIE, E. W. & RATNOFF, O. D. 1964. Waterfall Sequence for Intrinsic Blood Clotting. *Science*, 145, 1310-2.

DENT, J. A., BERKOWITZ, S. D., WARE, J., KASPER, C. K. & RUGGERI, Z. M. 1990. Identification of a cleavage site directing the immunochemical detection of molecular abnormalities in type IIA von Willebrand factor. *Proc Natl Acad Sci U S A*, 87, 6306-10.

EIKENBOOM, J. C., CASTAMAN, G., VOS, H. L., BERTINA, R. M. & RODEGHIERO, F. 1998. Characterization of the genetic defects in recessive type 1 and type 3 von Willebrand disease patients of Italian origin. *Thromb Haemost*, 79, 709-17.

EIKENBOOM, J. C., PLOOS VAN AMSTEL, H. K., REITSMA, P. H. & BRIET, E. 1992. Mutations in severe, type III von Willebrand's disease in the Dutch population: candidate missense and nonsense mutations associated with reduced levels of von Willebrand factor messenger RNA. *Thromb Haemost*, 68, 448-54.

EIKENBOOM, J. C., MATSUSHITA, T., REITSMA, P. H., TULEY, E. A., CASTAMAN, G., BRIET, E. & SADLER, J. E. 1996. Dominant type 1 von Willebrand disease caused by mutated cysteine residues in the D3 domain of von Willebrand factor. *Blood*, 88, 2433-41.

EIKENBOOM, J. C., REITSMA, P. H. & BRIET, E. 1994. Seeming homozygosity in type-IIb von Willebrand's disease due to a polymorphism within the sequence of a commonly used primer. *Ann Hematol*, 68, 139-41.

EIKENBOOM, J. C., REITSMA, P. H., PEERLINCK, K. M. & BRIET, E. 1993. Recessive inheritance of von Willebrand's disease type I. *Lancet*, 341, 982-6.

EL-MAARRI, O., HERBINIAUX, U., GRAW, J., SCHRODER, J., TERZIC, A., WATZKA, M., BRACKMANN, H. H., SCHRAMM, W., HANFLAND, P., SCHWAAB, R., MULLER, C. R. & OLDENBURG, J. 2005. Analysis of mRNA in hemophilia A patients with undetectable mutations reveals normal splicing in the factor VIII gene. *J Thromb Haemost*, 3, 332-9.

EL-MAARRI, O., SINGER, H., KLEIN, C., WATZKA, M., HERBINIAUX, U., BRACKMANN, H. H., SCHRODER, J., GRAW, J., MULLER, C. R., SCHRAMM, W., SCHWAAB, R., HAAF, T., HANFLAND, P. & OLDENBURG, J. 2006. Lack of F8 mRNA: a novel mechanism leading to hemophilia A. *Blood*, 107, 2759-65.

ENSEMBL GENOME BROWSER for *F8* and *VWF* gene transcripts
<http://www.ensembl.org>. Accessed 01 October 2012.

FAVALORO, E. J. 2000. Collagen binding assay for von Willebrand factor (VWF:CBA): detection of von Willebrands Disease (VWD), and discrimination of VWD subtypes, depends on collagen source. *Thromb Haemost*, 83, 127-35.

FAVALORO, E. J., LILLICRAP, D., LAZZARI, M. A., CATTANEO, M., MAZURIER, C., WOODS, A., MESCHENGIESER, S., BLANCO, A., KEMPFER, A. C., HUBBARD, A. & CHANG, A. 2004. von Willebrand disease: laboratory aspects of diagnosis and treatment. *Haemophilia*, 10 Suppl 4, 164-8.

FEDERICI, A. B. 1998. Diagnosis of von Willebrand disease. *Haemophilia*, 4, 654-60.

FEDERICI, A. B. 2003. The factor VIII/von Willebrand factor complex: basic and clinical issues. *Haematologica*, 88, EREP02.

FEDERICI, A. B. 2004. Clinical diagnosis of von Willebrand disease. *Haemophilia*, 10 Suppl 4, 169-76.

FEDERICI, A. B. 2009. Classification of inherited von Willebrand disease and implications in clinical practice. *Thromb Res*, 124 Suppl 1, S2-6.

GILL, J. C., ENDRES-BROOKS, J., BAUER, P. J., MARKS, W. J., JR. & MONTGOMERY, R. R. 1987. The effect of ABO blood group on the diagnosis of von Willebrand disease. *Blood*, 69, 1691-5.

GINSBURG, D. 1999. Molecular genetics of von Willebrand disease. *Thromb Haemost*, 82, 585-91.

GINSBURG, D. & BOWIE, E. J. 1992. Molecular genetics of von Willebrand disease. *Blood*, 79, 2507-19.

GINSBURG, D., HANDIN, R. I., BONTHRON, D. T., DONLON, T. A., BRUNS, G. A., LATT, S. A. & ORKIN, S. H. 1985. Human von Willebrand factor (vWF): isolation of complementary DNA (cDNA) clones and chromosomal localization. *Science*, 228, 1401-6.

GINSBURG, D. & SADLER, J. E. 1993. von Willebrand disease: a database of point mutations, insertions, and deletions. For the Consortium on von Willebrand Factor Mutations and Polymorphisms, and the Subcommittee on von Willebrand Factor of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis. *Thromb Haemost*, 69, 177-84.

GITSCHIER, J. 1988. Maternal duplication associated with gene deletion in sporadic hemophilia. *Am J Hum Genet*, 43, 274-9.

GITSCHIER, J., WOOD, W. I., GORALKA, T. M., WION, K. L., CHEN, E. Y., EATON, D. H., VEHAR, G. A., CAPON, D. J. & LAWN, R. M. 1984. Characterization of the human factor VIII gene. *Nature*, 312, 326-30.

GOODEVE, A. 2003. The incidence of inhibitor development according to specific mutations--and treatment? *Blood Coagul Fibrinolysis*, 14 Suppl 1, S17-21.

GOODEVE, A. 2007. Genetics of type 1 von Willebrand disease. *Curr Opin Hematol*, 14, 444-9.

GRAHAM, J. B., RIZZA, C. R., CHEDIAK, J., MANNUCCI, P. M., BRIET, E., LJUNG, R., KASPER, C. K., ESSIEN, E. M. & GREEN, P. P. 1986. Carrier detection in hemophilia A: a cooperative international study. I. The carrier phenotype. *Blood*, 67, 1554-9.

GUPTA, P. K., ADAMTZIKI, E., BUDDE, U., JAIPRAKASH, M., KUMAR, H., HARBECK-SEU, A., KANNAN, M., OYEN, F., OBSER, T., WEDEKIND, I., SAXENA, R. & SCHNEPPENHEIM, R. 2005. Gene conversions are a common cause of von Willebrand disease. *Br J Haematol*, 130, 752-8.

HABART, D., KALABOVA, D., NOVOTNY, M. & VORLOVA, Z. 2003. Thirty-four novel mutations detected in factor VIII gene by multiplex CSGE: modeling of 13 novel amino acid substitutions. *J Thromb Haemost*, 1, 773-81.

HABERICHTER, S. L., CASTAMAN, G., BUDDE, U., PEAKE, I., GOODEVE, A., RODEGHIERO, F., FEDERICI, A. B., BATLLE, J., MEYER, D., MAZURIER, C., GOUDEMAND, J., EIKENBOOM, J., SCHNEPPENHEIM, R., INGERSLEV, J., VORLOVA, Z., HABART, D., HOLMBERG, L., LETHAGEN, S., PASI, J., HILL, F. G. & MONTGOMERY, R. R. 2008. Identification of type 1 von Willebrand disease patients with reduced von Willebrand factor survival by assay of the VWF propeptide in the European study: molecular and clinical markers for the diagnosis and management of type 1 VWD (MCMDM-1VWD). *Blood*, 111, 4979-85.

HALLDEN, C., KNOBE, K. E., SJORIN, E., NILSSON, D. & LJUNG, R. 2012. Investigation of disease-associated factors in haemophilia A patients without detectable mutations. *Haemophilia*, 18, e132-7.

HAMSTeRS (HADB), the Haemophilia A Mutation, Structure, Test and Resource Site. <http://hadb.org.uk/>. Accessed 02 December 2012.

HASHEMI SOTEH, M., PEAKE, I. R., MARSDEN, L., ANSON, J., BATLLE, J., MEYER, D., FRESSINAUD, E., MAZURIER, C., GOUDEMAND, J., EIKENBOOM, J. & GOODEVE, A. 2007. Mutational analysis of the von Willebrand factor gene in type 1 von Willebrand disease using conformation sensitive gel electrophoresis: a comparison of fluorescent and manual techniques. *Haematologica*, 92, 550-3.

HICKSON, N., HAMPSHIRE, D., WINSHIP, P., GOUDEMAND, J., SCHNEPPENHEIM, R., BUDDE, U., CASTAMAN, G., RODEGHIRO, F., FEDERICI, A. B., JAMES, P., PEAKE, I., EIKENBOOM, J. & GOODEVE, A. 2010. von Willebrand factor variant p.Arg924Gln marks an allele associated with reduced von Willebrand factor and factor VIII levels. *J Thromb Haemost*, 8, 1986-93.

HILBERT, L., JORIEUX, S., FONTENAY-ROUPIE, M., GUICHETEAU, M., FRESSINAUD, E., MEYER, D. & MAZURIER, C. 2004. Expression of two type 2N von Willebrand disease mutations identified in exon 18 of von Willebrand factor gene. *Br J Haematol*, 127, 184-9.

HOFFBRAND, A. V., CATOVSKY D., AND TUDDENHAM, E.G.D. 2005. *Postgraduate Haematology*, Blackwell Publishing Ltd.

HOFFMAN, M. 2003. A cell-based model of coagulation and the role of factor VIIa. *Blood Rev*, 17 Suppl 1, S1-5.

HOLLESTELLE, M. J., THINNES, T., CRAIN, K., STIKO, A., KRUIJT, J. K., VAN BERKEL, T. J., LOSKUTOFF, D. J. & VAN MOURIK, J. A. 2001. Tissue distribution of factor VIII gene expression in vivo--a closer look. *Thromb Haemost*, 86, 855-61.

HOLMBERG, L. & NILSSON, I. M. 1985. von Willebrand disease. *Clin Haematol*, 14, 461-88.

HOLMBERG, L. & NILSSON, I. M. 1992. von Willebrand's disease. *Eur J Haematol*, 48, 127-41.

HU, X. Y., RAY, P. N. & WORTON, R. G. 1991. Mechanisms of tandem duplication in the Duchenne muscular dystrophy gene include both homologous and nonhomologous intrachromosomal recombination. *EMBO J*, 10, 2471-7.

International Society on Thrombosis and Hemostasis. VWF database. <http://www.vwf.group.shef.co.uk/>. Accessed November 30, 2012.

JACQUEMIN, M., DE MAEYER, M., D'OIRON, R., LAVEND'HOMME, R., PEERLINCK, K. & SAINT-REMY, J. M. 2003. Molecular mechanisms of mild and moderate hemophilia A. *J Thromb Haemost*, 1, 456-63.

JAFFE, E. A., HOYER, L. W. & NACHMAN, R. L. 1973. Synthesis of antihemophilic factor antigen by cultured human endothelial cells. *J Clin Invest*, 52, 2757-64.

JAMES, P. D., NOTLEY, C., HEGADORN, C., LEGGO, J., TUTTLE, A., TINLIN, S., BROWN, C., ANDREWS, C., LABELLE, A., CHIRINIAN, Y., O'BRIEN, L., OTHMAN, M., RIVARD, G., RAPSON, D., HOUGH, C. & LILLICRAP, D. 2007. The mutational spectrum of type 1 von Willebrand disease: Results from a Canadian cohort study. *Blood*, 109, 145-54.

- JAYANDHARAN, G., SHAJI, R. V., BAIDYA, S., NAIR, S. C., CHANDY, M. & SRIVASTAVA, A. 2005. Identification of factor VIII gene mutations in 101 patients with haemophilia A: mutation analysis by inversion screening and multiplex PCR and CSGE and molecular modelling of 10 novel missense substitutions. *Haemophilia*, 11, 481-91.
- JOHANSSON, A. M., HALLDEN, C., SALL, T. & LETHAGEN, S. 2011. Variation in the VWF gene in Swedish patients with type 1 von Willebrand Disease. *Ann Hum Genet*, 75, 447-55.
- JOURNET, A. M., SAFFARIPOUR, S. & WAGNER, D. D. 1993. Requirement for both D domains of the propolypeptide in von Willebrand factor multimerization and storage. *Thromb Haemost*, 70, 1053-7.
- JURLANDER, B., THIM, L., KLAUSEN, N. K., PERSSON, E., KJALKE, M., REXEN, P., JORGENSEN, T. B., OSTERGAARD, P. B., ERHARDTSEN, E. & BJORN, S. E. 2001. Recombinant activated factor VII (rFVIIa): characterization, manufacturing, and clinical development. *Semin Thromb Hemost*, 27, 373-84.
- KAZAZIAN, H. H., JR., WONG, C., YOUSSEF, H., SCOTT, A. F., PHILLIPS, D. G. & ANTONARAKIS, S. E. 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature*, 332, 164-6.
- KEENEY, S., BOWEN, D., CUMMING, A., ENAYAT, S., GOODEVE, A. & HILL, M. 2008. The molecular analysis of von Willebrand disease: a guideline from the UK Haemophilia Centre Doctors' Organisation Haemophilia Genetics Laboratory Network. *Haemophilia*, 14, 1099-111.
- KEENEY, S. & CUMMING, A. M. 2001. The molecular biology of von Willebrand disease. *Clin Lab Haematol*, 23, 209-30.
- KEENEY, S., MITCHELL, M. & GOODEVE, A. 2005. The molecular analysis of haemophilia A: a guideline from the UK haemophilia centre doctors' organization haemophilia genetics laboratory network. *Haemophilia*, 11, 387-97.
- KEENEY, S., NEVILLE-RUTHERFORD C., CUMMING A.M., BOLTON-MAGGS P., HAY C. 2007. Detection of partial F8 gene deletions in Haemophilia A carrier diagnosis. *Journal of Thrombosis and Haemostasis* (5), Supplement 2: abstract number O-S-014
- KEIGHTLEY, A. M., LAM, Y. M., BRADY, J. N., CAMERON, C. L. & LILLICRAP, D. 1999. Variation at the von Willebrand factor (vWF) gene locus is associated with plasma vWF:Ag levels: identification of three novel single nucleotide polymorphisms in the vWF gene promoter. *Blood*, 93, 4277-83.
- KENTER, A. L. & BIRSHTEN, B. K. 1981. Chi, a promoter of generalized recombination in lambda phage, is present in immunoglobulin genes. *Nature*, 293, 402-4.

- KIM, H. J., CHUNG, H. S., KIM, S. K., YOO, K. Y., JUNG, S. Y., PARK, I. A., LEE, K. O. & KIM, S. H. 2012. Mutation spectrum and inhibitor risk in 100 Korean patients with severe haemophilia A. *Haemophilia*.
- KLOPP, N., OLDENBURG, J., UEN, C., SCHNEPPENHEIM, R. & GRAW, J. 2002. 11 hemophilia A patients without mutations in the factor VIII encoding gene. *Thromb Haemost*, 88, 357-60.
- KOPPELMAN, S. J., VAN HOEIJ, M., VINK, T., LANKHOF, H., SCHIPHORST, M. E., DAMAS, C., VLOT, A. J., WISE, R., BOUMA, B. N. & SIXMA, J. J. 1996. Requirements of von Willebrand factor to protect factor VIII from inactivation by activated protein C. *Blood*, 87, 2292-300.
- LAFFAN, M., BROWN, S. A., COLLINS, P. W., CUMMING, A. M., HILL, F. G., KEELING, D., PEAKE, I. R. & PASI, K. J. 2004. The diagnosis of von Willebrand disease: a guideline from the UK Haemophilia Centre Doctors' Organization. *Haemophilia*, 10, 199-217.
- LAKICH, D., KAZAZIAN, H. H., JR., ANTONARAKIS, S. E. & GITSCHIER, J. 1993. Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nat Genet*, 5, 236-41.
- LANNOY, N., ABINET, I., DAHAN, K. & HERMANS, C. 2009. Identification of de novo deletion in the factor VIII gene by MLPA technique in two girls with isolated factor VIII deficiency. *Haemophilia*, 15, 797-801.
- LAVERGNE, J. M., BAHNAK, B. R., VIDAUD, M., LAURIAN, Y. & MEYER, D. 1992. A directed search for mutations in hemophilia A using restriction enzyme analysis and denaturing gradient gel electrophoresis. A study of seven exons in the factor VIII gene of 170 cases. *Nouv Rev Fr Hematol*, 34, 85-91.
- LENTING, P. J., VAN MOURIK, J. A. & MERTENS, K. 1998. The life cycle of coagulation factor VIII in view of its structure and function. *Blood*, 92, 3983-96.
- LETHAGEN, S., ISAKSSON, C., SCHAEDEL, C. & HOLMBERG, L. 2002. von Willebrand's disease caused by compound heterozygosity for a substitution mutation (T1156M) in the D3 domain of the von Willebrand factor and a stop mutation (Q2470X). *Thromb Haemost*, 88, 421-6.
- LEUER, M., OLDENBURG, J., LAVERGNE, J. M., LUDWIG, M., FREGIN, A., EIGEL, A., LJUNG, R., GOODEVE, A., PEAKE, I. & OLEK, K. 2001. Somatic mosaicism in hemophilia A: a fairly common event. *Am J Hum Genet*, 69, 75-87.
- LEVY, G. & GINSBURG, D. 2001. Getting at the variable expressivity of von Willebrand disease. *Thromb Haemost*, 86, 144-8.

- LILLICRAP, D., MURRAY, E. W., BENFORD, K., BLANCHETTE, V. S., RIVARD, G. E., WENSLEY, R. & GILES, A. R. 1991. Recurring mutations at CpG dinucleotides in the region of the von Willebrand factor gene encoding the glycoprotein Ib binding domain, in patients with type IIB von Willebrand's disease. *Br J Haematol*, 79, 612-7.
- LILLICRAP, D. P., WHITE, B. N., HOLDEN, J. J. & GILES, A. R. 1987. Carrier detection in the hemophilias. *Am J Hematol*, 26, 285-96.
- LIU, M. L., NAKAYA, S. & THOMPSON, A. R. 2002. Non-inversion factor VIII mutations in 80 hemophilia A families including 24 with alloimmune responses. *Thromb Haemost*, 87, 273-6.
- LIU, Q., NOZARI, G. & SOMMER, S. S. 1998. Single-tube polymerase chain reaction for rapid diagnosis of the inversion hotspot of mutation in hemophilia A. *Blood*, 92, 1458-9.
- LYONS, S. E., BRUCK, M. E., BOWIE, E. J. & GINSBURG, D. 1992. Impaired intracellular transport produced by a subset of type IIA von Willebrand disease mutations. *J Biol Chem*, 267, 4424-30.
- MACFARLANE, D. E., STIBBE, J., KIRBY, E. P., ZUCKER, M. B., GRANT, R. A. & MCPHERSON, J. 1975. Letter: A method for assaying von Willebrand factor (ristocetin cofactor). *Thromb Diath Haemorrh*, 34, 306-8.
- MACFARLANE, R. G. 1964. An Enzyme Cascade in the Blood Clotting Mechanism, and Its Function as a Biochemical Amplifier. *Nature*, 202, 498-9.
- MANCUSO, D. J., TULEY, E. A., WESTFIELD, L. A., LESTER-MANCUSO, T. L., LE BEAU, M. M., SORACE, J. M. & SADLER, J. E. 1991. Human von Willebrand factor gene and pseudogene: structural analysis and differentiation by polymerase chain reaction. *Biochemistry*, 30, 253-69.
- MANN, K. G. 1999. Biochemistry and physiology of blood coagulation. *Thromb Haemost*, 82, 165-74.
- MANNUCCI, P. M. 1998. Treatment of von Willebrand disease. *Int J Clin Lab Res*, 28, 211-4.
- MANNUCCI, P. M., BLOOM, A. L., LARRIEU, M. J., NILSSON, I. M. & WEST, R. R. 1984. Atherosclerosis and von Willebrand factor. I. Prevalence of severe von Willebrand's disease in western Europe and Israel. *Br J Haematol*, 57, 163-9.
- MANNUCCI, P. M., TAMARO, G., NARCHI, G., CANDOTTI, G., FEDERICI, A., ALTIERI, D. & TEDESCO, F. 1987. Life-threatening reaction to factor VIII concentrate in a patient with severe von Willebrand disease and alloantibodies to von Willebrand factor. *Eur J Haematol*, 39, 467-70.

MARGAGLIONE, M., CASTAMAN, G., MORFINI, M., ROCINO, A., SANTAGOSTINO, E., TAGARIELLO, G., TAGLIAFERRI, A. R., ZANON, E., BIOCCHI, M. P., CASTALDO, G., PEYVANDI, F., SANTACROCE, R., TORRICELLI, F., GRANDONE, E. & MANNUCCI, P. M. 2008. The Italian AICE-Genetics hemophilia A database: results and correlation with clinical phenotype. *Haematologica*, 93, 722-8.

MEYER, D., FRESSINAUD, E., GAUCHER, C., LAVERGNE, J. M., HILBERT, L., RIBBA, A. S., JORIEUX, S. & MAZURIER, C. 1997. Gene defects in 150 unrelated French cases with type 2 von Willebrand disease: from the patient to the gene. INSERM Network on Molecular Abnormalities in von Willebrand Disease. *Thromb Haemost*, 78, 451-6.

MICHAUX, G., ABBITT, K. B., COLLINSON, L. M., HABERICHTER, S. L., NORMAN, K. E. & CUTLER, D. F. 2006. The physiological function of von Willebrand's factor depends on its tubular storage in endothelial Weibel-Palade bodies. *Dev Cell*, 10, 223-32.

MILLAR, D. S., STEINBRECHER, R. A., WIELAND, K., GRUNDY, C. B., MARTINOWITZ, U., KRAWCZAK, M., ZOLL, B., WHITMORE, D., STEPHENSON, J., MIBASHAN, R. S. & ET AL. 1990. The molecular genetic analysis of haemophilia A; characterization of six partial deletions in the factor VIII gene. *Hum Genet*, 86, 219-27.

MILLER, C. H., BENSON, J., ELLINGSEN, D., DRIGGERS, J., PAYNE, A., KELLY, F. M., SOUCIE, J. M. & CRAIG HOOPER, W. 2012. F8 and F9 mutations in US haemophilia patients: correlation with history of inhibitor and race/ethnicity. *Haemophilia*, 18, 375-82.

MOHL, A., MARSCHALEK, R., MASSZI, T., NAGY, E., OBSER, T., OYEN, F., SALLAI, K., BODO, I. & SCHNEPPENHEIM, R. 2008. An Alu-mediated novel large deletion is the most frequent cause of type 3 von Willebrand disease in Hungary. *J Thromb Haemost*, 6, 1729-35.

MOHLKE, K. L., NICHOLS, W. C., REHEMTULLA, A., KAUFMAN, R. J., FAGERSTROM, H. M., RITVANEN, K. L., KEKOMAKI, R. & GINSBURG, D. 1996. A common frameshift mutation in von Willebrand factor does not alter mRNA stability but interferes with normal propeptide processing. *Br J Haematol*, 95, 184-91.

MRC-Holland. MRC-Holland MLPA homepage.

<http://www.mlpa.com/pages/indexpag.html>. Accessed 22 July 2012

MURRU, S., CASULA, L., PECORARA, M., MORI, P., CAO, A. & PIRASTU, M. 1990. Illegitimate recombination produced a duplication within the FVIII gene in a patient with mild hemophilia A. *Genomics*, 7, 115-8.

NACHMAN, R., LEVINE, R. & JAFFE, E. A. 1977. Synthesis of factor VIII antigen by cultured guinea pig megakaryocytes. *J Clin Invest*, 60, 914-21.

- NAKAYA, S. M., HSU, T. C., GERAGHTY, S. J., MANCO-JOHNSON, M. J. & THOMPSON, A. R. 2004. Severe hemophilia A due to a 1.3 kb factor VIII gene deletion including exon 24: homologous recombination between 41 bp within an Alu repeat sequence in introns 23 and 24. *J Thromb Haemost*, 2, 1941-5.
- NAYLOR, J., BRINKE, A., HASOCK, S., GREEN, P. M. & GIANNELLI, F. 1993. Characteristic mRNA abnormality found in half the patients with severe haemophilia A is due to large DNA inversions. *Hum Mol Genet*, 2, 1773-8.
- NESHEIM, M., PITTMAN, D. D., GILES, A. R., FASS, D. N., WANG, J. H., SLONOSKY, D. & KAUFMAN, R. J. 1991. The effect of plasma von Willebrand factor on the binding of human factor VIII to thrombin-activated human platelets. *J Biol Chem*, 266, 17815-20.
- NICHOLS, W. C. & GINSBURG, D. 1997. von Willebrand disease. *Medicine (Baltimore)*, 76, 1-20.
- NITU-WHALLEY, I. C., LEE, C. A., BROWN, S. A., RIDDELL, A. & HERMANS, C. 2003. The role of the platelet function analyser (PFA-100) in the characterization of patients with von Willebrand's disease and its relationships with von Willebrand factor and the ABO blood group. *Haemophilia*, 9, 298-302.
- O'BRIEN, L. A., JAMES, P. D., OTHMAN, M., BERBER, E., CAMERON, C., NOTLEY, C. R., HEGADORN, C. A., SUTHERLAND, J. J., HOUGH, C., RIVARD, G. E., O'SHAUNESSEY, D. & LILLICRAP, D. 2003. Founder von Willebrand factor haplotype associated with type 1 von Willebrand disease. *Blood*, 102, 549-57.
- O'BRIEN, L. A., SUTHERLAND, J. J., WEAVER, D. F. & LILLICRAP, D. 2005. Theoretical structural explanation for Group I and Group II, type 2A von Willebrand disease mutations. *J Thromb Haemost*, 3, 796-7.
- OLDENBURG, J., EL-MAARRI, O. & SCHWAAB, R. 2002. Inhibitor development in correlation to factor VIII genotypes. *Haemophilia*, 8 Suppl 2, 23-9.
- OLDENBURG, J., IVASKEVICIUS, V., ROST, S., FREGIN, A., WHITE, K., HOLINSKI-FEDER, E., MULLER, C. R. & WEBER, B. H. 2001. Evaluation of DHPLC in the analysis of hemophilia A. *J Biochem Biophys Methods*, 47, 39-51.
- OLDENBURG, J., SCHRODER, J., BRACKMANN, H. H., MULLER-REIBLE, C., SCHWAAB, R. & TUDDENHAM, E. 2004. Environmental and genetic factors influencing inhibitor development. *Semin Hematol*, 41, 82-8.
- PADILLA, A., MOAKE, J. L., BERNARDO, A., BALL, C., WANG, Y., ARYA, M., NOLASCO, L., TURNER, N., BERNDT, M. C., ANVARI, B., LOPEZ, J. A. & DONG, J. F. 2004. P-selectin anchors newly released ultralarge von Willebrand factor multimers to the endothelial cell surface. *Blood*, 103, 2150-6.

PAVLOVA, A., FORSTER, T., DELEV, D., SCHRODER, J., EL-MAARRI, O., MULLER-REIBLE, C. & OLDENBURG, J. 2008. Heterozygous large deletions of Factor 8 gene in females identified by multiplex PCR-LC. *Haemophilia*, 14, 599-606.

PEAKE, I. R., LIDDELL, M. B., MOODIE, P., STANDEN, G., MANCUSO, D. J., TULEY, E. A., WESTFIELD, L. A., SORACE, J. M., SADLER, J. E., VERWEIJ, C. L. & ET AL. 1990. Severe type III von Willebrand's disease caused by deletion of exon 42 of the von Willebrand factor gene: family studies that identify carriers of the condition and a compound heterozygous individual. *Blood*, 75, 654-61.

PEYVANDI, F., JAYANDHARAN, G., CHANDY, M., SRIVASTAVA, A., NAKAYA, S. M., JOHNSON, M. J., THOMPSON, A. R., GOODEVE, A., GARAGIOLA, I., LAVORETANO, S., MENEGATTI, M., PALLA, R., SPREAFICO, M., TAGLIABUE, L., ASSELTA, R., DUGA, S. & MANNUCCI, P. M. 2006. Genetic diagnosis of haemophilia and other inherited bleeding disorders. *Haemophilia*, 12 Suppl 3, 82-9.

POSAN, E., MCBANE, R. D., GRILL, D. E., MOTSKO, C. L. & NICHOLS, W. L. 2003. Comparison of PFA-100 testing and bleeding time for detecting platelet hypofunction and von Willebrand disease in clinical practice. *Thromb Haemost*, 90, 483-90.

RASCHE, H. (2001). "Haemostasis and thrombosis: an overview." *Eur Heart J Supplements* 3 (Suppl Q): Q3-Q7.

RAFATI, M., RAVANBOD, S., HOSEINI, A., RASSOULZADEGAN, M., JAZEBI, M., ENAYAT, M. S., ALA, F. A. & GHAFARI, S. R. 2011. Identification of ten large deletions and one duplication in the F8 gene of eleven unrelated Iranian severe haemophilia A families using the multiplex ligation-dependent probe amplification technique. *Haemophilia*, 17, 705-7.

ROBERTSON, J., LILLICRAP, D. & JAMES, P. D. 2008. von Willebrand disease. *Pediatr Clin North Am*, 55, 377-92, viii-ix.

RODEGHIRO, F., CASTAMAN, G. & DINI, E. 1987. Epidemiological investigation of the prevalence of von Willebrand's disease. *Blood*, 69, 454-9.

ROSSETTI, L. C., RADIC, C. P., LARRIPA, I. B. & DE BRASI, C. D. 2005. Genotyping the hemophilia inversion hotspot by use of inverse PCR. *Clin Chem*, 51, 1154-8.

ROSSETTI, L. C., RADIC, C. P., LARRIPA, I. B. & DE BRASI, C. D. 2008. Developing a new generation of tests for genotyping hemophilia-causative rearrangements involving int22h and int1h hotspots in the factor VIII gene. *J Thromb Haemost*, 6, 830-6.

- ROST, S., LOFFLER, S., PAVLOVA, A., MULLER, C. R. & OLDENBURG, J. 2008. Detection of large duplications within the factor VIII gene by MLPA. *J Thromb Haemost*, 6, 1996-9.
- RUGGERI, Z. M. 1997. von Willebrand factor. *J Clin Invest*, 100, S41-6.
- RUGGERI, Z. M. 2003. von Willebrand factor. *Curr Opin Hematol*, 10, 142-9.
- RUGGERI, Z. M., MANNUCCI, P. M., LOMBARDI, R., FEDERICI, A. B. & ZIMMERMAN, T. S. 1982. Multimeric composition of factor VIII/von Willebrand factor following administration of DDAVP: implications for pathophysiology and therapy of von Willebrand's disease subtypes. *Blood*, 59, 1272-8.
- SADLER, J. E. 1994. A revised classification of von Willebrand disease. For the Subcommittee on von Willebrand Factor of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis. *Thromb Haemost*, 71, 520-5.
- SADLER, J. E., BUDDE, U., EIKENBOOM, J. C., FAVALORO, E. J., HILL, F. G., HOLMBERG, L., INGERSLEV, J., LEE, C. A., LILLICRAP, D., MANNUCCI, P. M., MAZURIER, C., MEYER, D., NICHOLS, W. L., NISHINO, M., PEAKE, I. R., RODEGHIERO, F., SCHNEPPENHEIM, R., RUGGERI, Z. M., SRIVASTAVA, A., MONTGOMERY, R. R. & FEDERICI, A. B. 2006. Update on the pathophysiology and classification of von Willebrand disease: a report of the Subcommittee on von Willebrand Factor. *J Thromb Haemost*, 4, 2103-14.
- SADLER, J. E. 2009. von Willebrand factor assembly and secretion. *J Thromb Haemost*, 7 Suppl 1, 24-7.
- SADLER, J. E., MANNUCCI, P. M., BERNTORP, E., BOCHKOV, N., BOULYJENKOV, V., GINSBURG, D., MEYER, D., PEAKE, I., RODEGHIERO, F. & SRIVASTAVA, A. 2000. Impact, diagnosis and treatment of von Willebrand disease. *Thromb Haemost*, 84, 160-74.
- SADLER, J. E. & RODEGHIERO, F. 2005. Provisional criteria for the diagnosis of VWD type 1. *J Thromb Haemost*, 3, 775-7.
- SAIKI, R. K., SCHARF, S., FALOONA, F., MULLIS, K. B., HORN, G. T., ERLICH, H. A. & ARNHEIM, N. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, 230, 1350-4.
- SAIKI, R. K., GELFAND, D. H., STOFFEL, S., SCHARF, S. J., HIGUCHI, R., HORN, G. T., MULLIS, K. B. & ERLICH, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239, 487-91.
- SAVAGE, B., SALDIVAR, E. & RUGGERI, Z. M. 1996. Initiation of platelet adhesion by arrest onto fibrinogen or translocation on von Willebrand factor. *Cell*, 84, 289-97.

SCHNEPPENHEIM, R., BUDDE, U., KREY, S., DREWKE, E., BERGMANN, F., LECHLER, E., OLDENBURG, J. & SCHWAAB, R. 1996. Results of a screening for von Willebrand disease type 2N in patients with suspected haemophilia A or von Willebrand disease type 1. *Thromb Haemost*, 76, 598-602.

SCHNEPPENHEIM, R., BUDDE, U., OBSER, T., BRASSARD, J., MAINUSCH, K., RUGGERI, Z. M., SCHNEPPENHEIM, S., SCHWAAB, R. & OLDENBURG, J. 2001. Expression and characterization of von Willebrand factor dimerization defects in different types of von Willebrand disease. *Blood*, 97, 2059-66.

SCHNEPPENHEIM, R., CASTAMAN, G., FEDERICI, A. B., KREUZ, W., MARSCHALEK, R., OLDENBURG, J., OYEN, F. & BUDDE, U. 2007. A common 253-kb deletion involving VWF and TMEM16B in German and Italian patients with severe von Willebrand disease type 3. *J Thromb Haemost*, 5, 722-8.

SCHNEPPENHEIM, R., KREY, S., BERGMANN, F., BOCK, D., BUDDE, U., LANGE, M., LINDE, R., MITTLER, U., MEILI, E., MERTES, G. & ET AL. 1994. Genetic heterogeneity of severe von Willebrand disease type III in the German population. *Hum Genet*, 94, 640-52.

SCHOUTEN, J. P., MCELGUNN, C. J., WAAIJER, R., ZWIJNENBURG, D., DIEPVENS, F. & PALS, G. 2002. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res*, 30, e57.

SCHWAAB, R., OLDENBURG, J., SCHWAAB, U., JOHNSON, D. J., SCHMIDT, W., OLEK, K., BRACKMAN, H. H. & TUDDENHAM, E. G. 1995. Characterization of mutations within the factor VIII gene of 73 unrelated mild and moderate haemophiliacs. *Br J Haematol*, 91, 458-64.

SURDHAR, G. K., ENAYAT, M. S., LAWSON, S., WILLIAMS, M. D. & HILL, F. G. 2001. Homozygous gene conversion in von Willebrand factor gene as a cause of type 3 von Willebrand disease and predisposition to inhibitor development. *Blood*, 98, 248-50.

SUTHERLAND, M. S., CUMMING, A. M., BOWMAN, M., BOLTON-MAGGS, P. H., BOWEN, D. J., COLLINS, P. W., HAY, C. R., WILL, A. M. & KEENEY, S. 2009. A novel deletion mutation is recurrent in von Willebrand disease types 1 and 3. *Blood*, 114, 1091-8.

SZTUKOWSKA, M., GALLINARO, L., CATTINI, M. G., PONTARA, E., SARTORELLO, F., DAIDONE, V., PADRINI, R., PAGNAN, A. & CASONATO, A. 2008. von Willebrand factor propeptide makes it easy to identify the shorter von Willebrand factor survival in patients with type 1 and type Vicenza von Willebrand disease. *Br J Haematol*, 143, 107-14.

- THEOPHILUS, B. D., BAUGH, L. M., GUILLIATT, A. M., MOTWANI, J. & WILLIAMS, M. D. 2011. Deletions and duplications in the factor VIII gene identified using multiplex ligation-dependent probe amplification. *J Thromb Haemost*, 9, 605-7.
- TJERNBERG, P., VOS, H. L., CASTAMAN, G., BERTINA, R. M. & EIKENBOOM, J. C. 2004. Dimerization and multimerization defects of von Willebrand factor due to mutated cysteine residues. *J Thromb Haemost*, 2, 257-65.
- VERWEIJ, C. L., DIERGAARDE, P. J., HART, M. & PANNEKOEK, H. 1986. Full-length von Willebrand factor (vWF) cDNA encodes a highly repetitive protein considerably larger than the mature vWF subunit. *EMBO J*, 5, 1839-47.
- VERWEIJ, C. L., HART, M. & PANNEKOEK, H. 1988. Proteolytic cleavage of the precursor of von Willebrand factor is not essential for multimer formation. *J Biol Chem*, 263, 7921-4.
- VON WILLEBRAND, E.A. (1926). "Hereditar pseudoheemofil." Finska Lakar. Hand., **68**, 87-112.
- VON WILLEBRAND, E.A. (1931). "Ueber hereditaere Pseudohaemophile. Acta Med. Scand., **76**, 521-550.
- VOORBERG, J., FONTIJN, R., CALAFAT, J., JANSSEN, H., VAN MOURIK, J. A. & PANNEKOEK, H. 1991. Assembly and routing of von Willebrand factor variants: the requirements for disulfide-linked dimerization reside within the carboxy-terminal 151 amino acids. *J Cell Biol*, 113, 195-205.
- VOORBERG, J., FONTIJN, R., VAN MOURIK, J. A. & PANNEKOEK, H. 1990. Domains involved in multimer assembly of von willebrand factor (vWF): multimerization is independent of dimerization. *EMBO J*, 9, 797-803.
- WHITE, G. C., 2ND, ROSENDAAL, F., ALEDORT, L. M., LUSHER, J. M., ROTHSCCHILD, C. & INGERSLEV, J. 2001. Definitions in hemophilia. Recommendation of the scientific subcommittee on factor VIII and factor IX of the scientific and standardization committee of the International Society on Thrombosis and Haemostasis. *Thromb Haemost*, 85, 560.
- WISE, R. J., PITTMAN, D. D., HANDIN, R. I., KAUFMAN, R. J. & ORKIN, S. H. 1988. The propeptide of von Willebrand factor independently mediates the assembly of von Willebrand multimers. *Cell*, 52, 229-36.
- WOODS-SAMUELS, P., KAZAZIAN, H. H., JR. & ANTONARAKIS, S. E. 1991. Nonhomologous recombination in the human genome: deletions in the human factor VIII gene. *Genomics*, 10, 94-101.

XIE, F., WANG, X., COOPER, D. N., CHUZHANOVA, N., FANG, Y., CAI, X., WANG, Z. & WANG, H. 2006. A novel Alu-mediated 61-kb deletion of the von Willebrand factor (VWF) gene whose breakpoints co-locate with putative matrix attachment regions. *Blood Cells Mol Dis*, 36, 385-91.

YADEGARI, H., DRIESEN, J., HASS, M., BUDDE, U., PAVLOVA, A. & OLDENBURG, J. 2011. Large deletions identified in patients with von Willebrand disease using multiple ligation-dependent probe amplification. *J Thromb Haemost*, 9, 1083-6.

YOUNG, M., INABA, H., HOYER, L. W., HIGUCHI, M., KAZAZIAN, H. H., JR. & ANTONARAKIS, S. E. 1997. Partial correction of a severe molecular defect in hemophilia A, because of errors during expression of the factor VIII gene. *Am J Hum Genet*, 60, 565-73.

YOUSSEF, H., ANTONARAKIS, S. E., ARONIS, S., TSIFTIS, G., PHILLIPS, D. G. & KAZAZIAN, H. H., JR. 1987. Characterization of five partial deletions of the factor VIII gene. *Proc Natl Acad Sci U S A*, 84, 3772-6.

ZHANG, Z. P., BLOMBACK, M., EGBERG, N., FALK, G. & ANVRET, M. 1994. Characterization of the von Willebrand factor gene (VWF) in von Willebrand disease type III patients from 24 families of Swedish and Finnish origin. *Genomics*, 21, 188-93.

ZHANG, Z. P., BLOMBACK, M., NYMAN, D. & ANVRET, M. 1993. Mutations of von Willebrand factor gene in families with von Willebrand disease in the Aland Islands. *Proc Natl Acad Sci U S A*, 90, 7937-40.

ZHANG, Z. P., FALK, G., BLOMBACK, M., EGBERG, N. & ANVRET, M. 1992. A single cytosine deletion in exon 18 of the von Willebrand factor gene is the most common mutation in Swedish vWD type III patients. *Hum Mol Genet*, 1, 767-8.

ZHANG, B. & GINSBURG, D. 2004. Familial multiple coagulation factor deficiencies: new biologic insight from rare genetic bleeding disorders. *J Thromb Haemost*, 2, 1564-72.

ZHANG, B., MCGEE, B., YAMAOKA, J. S., GUGLIELMONE, H., DOWNES, K. A., MINOLDO, S., JARCHUM, G., PEYVANDI, F., DE BOSCH, N. B., RUIZ-SAEZ, A., CHATELAIN, B., OLPINSKI, M., BOCKENSTEDT, P., SPERL, W., KAUFMAN, R. J., NICHOLS, W. C., TUDDENHAM, E. G. & GINSBURG, D. 2006. Combined deficiency of factor V and factor VIII is due to mutations in either LMAN1 or MCFD2. *Blood*, 107, 1903-7.

ZIMMERMANN, M. A., OLDENBURG, J., MULLER, C. R. & ROST, S. 2010. Characterization of duplication breakpoints in the factor VIII gene. *J Thromb Haemost*, 8, 2696-704.

APPENDICES

Appendix 1: SALSA MLPA P178-B1 *F8* Probemix

Length	SALSA MLPA probe	Chromosomal position
64-70-76-82	Q-fragments: DNA quantity; only visible with less than 100 ng sample DNA	
88-92-96	D-fragments: Low signal of 88 or 96 nt fragment indicates incomplete denaturation	
100	X-fragment: Specific for the X chromosome	
105	Y-fragment: Specific for the Y chromosome	
136	Reference probe 07103-L14743	Xp22.11
142	<i>F8</i> probe 05610-L05060	Exon 1
148	<i>F8</i> probe 05618-L0568	Exon 7
154	<i>F8</i> probe 05627-L06055	Exon 14
160	<i>F8</i> probe 05633-L05083	Exon 20
166	Reference probe 02927-L03721	Xq27.3
172	<i>F8</i> probe 05612-L05062	Exon 2
179	<i>F8</i> probe 05619-L14947	Exon 8
186	<i>F8</i> probe 06290-L17545	Exon 6
191	<i>F8</i> probe 05626-L17546	Exon 14
197	<i>F8</i> probe 05634-L17547	Exon 21
202	Reference probe 13203-L14524	Xq13.1
209	<i>F8</i> probe 05611-L17548	Exon 1
215	<i>F8</i> probe 05620-L17549	Exon 9
221	<i>F8</i> probe 06288-L05892	Exon 5
229	<i>F8</i> probe 05641-L05091	Exon 27
238	Reference probe 06910-L06490	Xq22.1
247	<i>F8</i> probe 05617-L05067	Exon 7
256	<i>F8</i> probe 05621-L05071	Exon 10
265	<i>F8</i> probe 05628-L05078	Exon 15
274	<i>F8</i> probe 07044-L05899	Exon 12
283	Reference probe 06035-L05490	Xq13.1
292	Reference probe 02920-L02314	Xp21.3
298	<i>F8</i> probe 05614-L05891	Exon 4
305	<i>F8</i> probe 06287-L06058	Exon 11
313	<i>F8</i> probe 06506-L06340	Exon 3
319	<i>F8</i> probe 05629-L05079	Exon 16
328	<i>F8</i> probe 05637-L05087	Exon 24
337	Reference probe 02932-L02323	Xq28
346	Reference probe 05125-L04515	Xq26.2
355	<i>F8</i> probe 05623-L05893	Exon 12
364	<i>F8</i> probe 05630-L05080	Exon 17
372	<i>F8</i> probe 14164-L15766	Exon 25
382	Reference probe 02908-L02302	Xq23
391	<i>F8</i> probe 07045-L05898	Exon 22
400	<i>F8</i> probe 05624-L05074	Exon 13
409	<i>F8</i> probe 05631-L05081	Exon 18
418	<i>F8</i> probe 13707-L15181	Exon 26

426	Reference probe 13207-L14528	Xp11.3
434	Reference probe 13118-L15558	Xq21.2
444	<i>F8</i> probe 05613-L14948	Exon 3
452	<i>F8</i> probe 05625-L14949	Exon 14
460	<i>F8</i> probe 05632-L14950	Exon 19
468	<i>F8</i> probe 13708-L15182	Exon 27
476	Reference probe 01391-L01039	Xp21.1

Appendix 2a: SALSA MLPA P011-B1 (Kit 1) VWF Probemix

Length	SALSA MLPA probe	Chromosomal position
64-70-76-82	Q-fragments: DNA quantity; only visible with less than 100 ng sample DNA	
88-92-96	D-fragments: Low signal of 88 or 96 nt fragment indicates incomplete denaturation	
100	X-fragment: Specific for the X chromosome	
105	Y-fragment: Specific for the Y chromosome	
124	Reference probe 09176-L09350	17q11
130	VWF probe 11314-L12040	Exon 5
136	VWF probe 11332-L12057	Exon 39
141	Reference probe 10301-L10813	11q14
148	VWF probe 11335-L12060	Exon 45
154	VWF probe 11325-L12050	Exon 24
160	VWF probe 11330-L12055	Exon 36
169	VWF probe 11321-L12046	Exon 14
178	VWF probe 11339-L12064	Exon 50
185	VWF probe 11328-L12053	Exon 33
195	Reference probe 10375-L12864	9q34
202	VWF probe 11320-L12045	Exon 13
208	VWF probe 11331-L13985	Exon 38
214	VWF probe 11319-L12044	Exon 9
226	VWF probe 11324-L12049	Exon 23
232	VWF probe 11338-L12063	Exon 49
238	Reference probe 10137-L10599	18q11
247	VWF probe 113226-L12051	Exon 26
256	VWF probe 11334-L14242	Exon 43
265	VWF probe 13422-L14877	Exon 2
274	VWF probe 11336-L12061	Exon 46
283	VWF probe 11329-L12054	Exon 34
301	Reference probe 09986-L10445	7q22
310	VWF probe 11316-L12042	Exon 6
319	VWF probe 11333-L12058	Exon 42
337	VWF probe 13423-L14878	Exon 2
346	VWF probe 11322-L12047	Exon 18
355	VWF probe 11318-L12043	Exon 8
364	VWF probe 11323-L12048	Exon 20
373	Reference probe 10718-L11300	6p12
383	VWF probe 11327-L12052	Exon 28
391	VWF probe 13426-L14881	Exon 47
400	Reference probe 08949-L09044	4q32
409	Reference probe 10053-L10477	8q22
418	VWF probe 11340-L12065	Exon 52
426	VWF probe 11315-L12041	Exon 6
434	Reference probe 10876-L11546	15q12

Appendix 2b: SALSA MLPA P012-B1 (Kit 2) VWF Probemix

Length	SALSA MLPA probe	Chromosomal position
64-70-76-82	Q-fragments: DNA quantity; only visible with less than 100 ng sample DNA	
88-92-96	D-fragments: Low signal of 88 or 96 nt fragment indicates incomplete denaturation	
100	X-fragment: Specific for the X chromosome	
105	Y-fragment: Specific for the Y chromosome	
124	Reference probe 09176-L09350	17q11
130	VWF probe 11346-L12071	Exon 12
136	VWF probe 11362-L12087	Exon 40
148	VWF probe 11355-L12080	Exon 28
154	VWF probe 11356-L12081	Exon 29
160	VWF probe 11359-L12084	Exon 32
172	VWF probe 11352-L12077	Exon 22
178	VWF probe 11354-L12079	Exon 27
184	VWF probe 11350-L12075	Exon 19
193	Reference probe 03861-L03244	3p24
202	VWF probe 11366-L12091	Exon 51
208	VWF probe 11363-L12088	Exon 41
215	VWF probe 11342-L12067	Exon 4
223	VWF probe 12799-L14243	Exon 37
232	VWF probe 11341-L12066	Exon 1
238	Reference probe 08951-L09046	4q32
247	VWF probe 11351-L12076	Exon 21
256	VWF probe 11349-L12074	Exon 17
265	Reference probe 08345-L08277	17q21
274	VWF probe 11358-L12083	Exon 31
283	VWF probe 11364-L12089	Exon 44
292	VWF probe 13425-L14880	Exon 47
301	Reference probe 09986-L10445	7q22
310	VWF probe 11344-L12069	Exon 10
322	VWF probe 11353-L12078	Exon 25
328	VWF probe 11348-L12073	Exon 16
337	Reference probe 10376-L10928	9q34
346	VWF probe 11360-L12085	Exon 35
355	VWF probe 11365-L12090	Exon 48
364	VWF probe 11347-L12072	Exon 15
373	Reference probe 06745-L06349	8q12
382	VWF probe 11343-L12068	Exon 7
391	VWF probe 11357-L12082	Exon 30
400	Reference probe 10174-L10922	18p11
409	Reference probe 09893-L10306	16p13
418	VWF probe 11345-L12070	Exon 11
427	VWF probe 06683-L06261	Exon 3
434	Reference probe 11665-L12432	14q24

Appendix 3: VWF sequence with probe locations

Partial sequence (20 nt adjacent to ligation site) of the VWF MLPA probes are indicated for the 52 exons of VWF.

Key

Exons – red text

MLPA probes adjacent to ligation site – highlighted in green and pink

E1

CTCAACTTCC	CTGGAGTCCA	GCCCTGGAAG	CTGGATCAGG	AAGCTGTGTT	GTTCTACTGT
GATTCCCCCT	GGCCTGTATC	AGCTTGCCCT	GAAACAACCA	GCATTCCCTGG	TTATCCCACA
CAGGTGGGGC	ACTCTAGGAA	GACCAGGGAT	CAAGTGTGGG	GGTGTAGGGA	TAGGGGGTGT
TTGGGGAGGG	CAAGGCAGTT	AATTAAGGCA	GCTGCCAGGA	GGTCTCCCTC	CAAACCTCTAC
AAAGCTTTAT	CAGCTTGGAG	GTACTTCTAA	TACCATTTCC	TTTCATTGTT	TCCTTTTGGT
AATTAAGAGG	AGGCCAATCC	CCTGTTGTGG	CAGCTCACAG	CTATTGTGGT	GGGAAAGGGA
GGGTGGTTGG	TGGATGTCAC	AGCTTGGGCT	TTATCTCCCC	CAGCAGTGGG	GACTCCACAG
CCCCTGGGCT	ACATAACAGC	AAGACAGTCC	GGAGCTGTAG	CAGACCTGAT	TGAGCCTTTG
CAGCAGCTGA	GAGCATGGCC	TAGGGTGGGC	GGCACCATTG	TCCA	CAGCT
GGGACCTTGG	AGATAGCCGC	AGCCCTCATT	TGCAGGGGAA	GGTATGGCCT	TTGGAAGGAG
AGCTGGCTCA	GTTGTGGGAG	GAAGATGCAG	GACTGACTGA	TCCCTGCTCC	TGGGGAGCTG
GAGTTCTCTG	TCGCTGGACT	AGAAGGGCTT	TGTTTGGAGG	GGCAATTCAA	TTCAGCCAGG

E2

TGAGGCCAG	CGGGCCTGGA	GGGAAGGCAG	CGTTTTCTCT	CTGTCTTGGG	CCTGGGACTG
CTGTCTGTGG	AAAGGTGCCC	ACAGGTCCCA	GCTCACAGCG	ATTGTTACCC	TTGGGCCTGG
CACTGGCCAG	GGGTTTTTTC	GGGGGCCAGA	AGTCCATGTT	CAAAGGGGAA	AAGGGGGTCA
CGAGGATCAA	TCTTTTCTCC	TGCTTTAAAG	AAATGTTTTT	GCTACTGCAT	GCCCTGATAG
TCGCCACACC	AGCAGCCGCC	TACCTGGGCA	GCAATGACCA	GCTCACGTCT	CTTGCTTCTT
TGCAGATGAT	TCCTGCCAGA	TTTGCCGGGG	TGCTGCTTGC	TCTGGCCCTC	ATTTTGCCAG
GTAGGTACAA	AAGGGCCTCC	ATTTCTCATT	CCTG	CCCCAG	GGCCATCTGG
TTCCGGGAAT	CAGCAGGTGT	GTCTGGAGCT	CACCTGTGTG	CCCAGCCCTA	ACTTAGGCTG

E3

CAGCCCAGTT	TCTATCAGAG	GCAGTCTAAT	TTCTACGAGG	CCAGAGAGGT	TTGAGCTGAT
GGTCCCAGTT	GTGCCCTGAG	ATCACCAGCC	CAACCTGTGG	CCTCTCCCTC	CAGGGACCTT
TTGTGCAGAA	GGAAGTCGCG	GCAGGTCATC	CACGGCCCGA	TGCAGCCTTT	TGGAACTGTA
CTCGTCAAC	ACCTTTGATG	GGAGCATGTA	CAGCTTTGCG	GGATACTGCA	GTTACCTCCT
GGCAGGGGGC	TGCCAGAAAC	GCTCCTTCTC	GATTATTGGT	GAGTTCTGGG	CACTGCAGGG
AGGACTTCAG	AGGGAGGGCT	GGCTGAGCTC	AGCCCTGGTG	TGGGGGAGGA	TTCTGTCTCT
CAGGACAGTG	TCTGAGTGGA	AAGGTCAGTG	CTGAGAACAA	GGAGAGGAAC	AGCCTTTCTG
TGACACGTAG	CCCCTCTTGG	CTTTCCCCGG	GTCTCTCCCC	ACGGGAGCCG	GGTGGGATGG
ATGAAGAGAG	TCTTCATCTT	TGGTAGTCCA	CTGTGTCCGT	TGCTCTGGGG	CCCGGCGATG
CCCTGGGAAC	TCCACAGCAT	CAAGGCAAAT	GATGAAGTAG	AGAAGGTGCT	TTGGAACGTG
TAAAACTCCT	TGCCAGAGAG	AAGACTCGTT	GTTGTTTTCT	TGGTGGCCTG	TGGATCAGAA

E4

TTTAGATGAA	TGGAAAAATG	GGTGAGAAAA	TGAGATTTTT	AGGGAGATTT	TTGTTGCTCG
CTTGTAAGA	CTTTTGGGG	CGTTTTCTGC	TGAGAAAAGG	TTACGTAGAT	AATGATTCTT
AATCAATGTA	TTCATTTTTT	GAGAGGAGTA	ATAATCACTA	CTATTGACTT	TTTTCTCTTT
CAGGGGACTT	CCAGAATGGC	AAGAGAGTGA	GCCTCTCCGT	GTATCTTCGG	GAATTCTTTC
ACATCCATT	TTGTTGTCAAT	GGTACCGTGA	CACAGGGGGA	CCAAAGGTAA	GCCAACAATG
TCTGAGTTAG	AAAGGACCCT	AGGGATCCCC	TGACACAACC	CCCTCATTTT	TAGATGAGGA
AGCTGGGGCC	CAGAGAATGG	AAGCAAATGT	TCCAAGGAAG	TGAGTAGCAG	GGCTGGGTGA

E5

CCTCAGCCAC	TTTGCACCAT	GTTCTGAACC	CTACAACATG	GGGTTGGGGT	TAGAAGGTGG
GAGAGACATC	CAGAAAATGC	ACAAGAAGCC	CACTTCTGAA	CTTAGCCTTT	GCCCTCCAGA
GTCTCCATGC	CCTATGCCTC	CAAAGGGCTG	TATCTAGAAA	CTGAGGCTGG	GTAATAAAG
CTGTCCGGTG	AGGCCTATGG	CTTTGTGGCC	AGGATCGATG	GCAGCCCAA	CTTTCAATGTC
CTGCTGTCTG	ACAGATACTT	CAACAAGACC	TGCGGGCTGT	GTGGCAACTT	TAACATCTTT
GCTGAAGATG	ACTTTATGAC	CCAAGAAGGT	AAGATGTTCT	GGGATACCAT	TTCCCTAAAG
TGTGGCCATG	CTTTTTATTT	CCTTGCTCAT	AAACCTTCAC	TAACATGCCT	TCCCTGGCAT
TCAAGCCTCA	CTGTGACCTC	ACCTCAATTT	ATGTTGCCAA	CCTTATCTCT	TTGCATTCCA

E6

AGAGATGGGT	CAAGTCAAAA	CTCCTGTGCT	GCTCAGTAGT	GGGATTGCAC	TTGTGAATAG
CCGCTGCACT	CCAGCCTGGG	AAACATAGCC	AGACCCTGTC	TCTTTATAAA	AAATTAATAAC
AAAACACACA	AAACCACCAG	CAGACCTAGA	ATTTTCACCC	AGAAGTTCTG	GGACAGGCAT
AACTGAAGCA	TTACTTTCCT	GAAACTTTCC	TCACAGGGA	CTTTGACCTC	GGACCTTTAT
GACTTTGCCA	ACTCATGGGC	TCTGAGCAGT	GGAACAACGT	GGTCTGAACG	GGCATCTCCT
CCCAGCAGCT	CATGCAACAT	CTCCTCTGGG	GAAATGCAGA	AGGTGGGTGT	GGACTGGCCT
GGGTGCACCT	GGATGGTGTG	TGATTTCTGG	ATCTAAAAGA	CAGAAGGACT	CAGTCTCATA
TCCTTCCATC	TGGGGGAGGA	ATGGACTTAC	GCAGGGCCAT	TTCTTCCAAA	ACTAACTGTG
GCTAGAGTCT	AATTCTAATA	CATCTCGAGC	CTGAAGCTCT	AAAAATGAGT	CTGGGCTAAT
GACTTGGTGT	AGTGTCTCAG	TCCATTTGTG	CTGCTATAAC	AAATGCCTGA	GACTGGGTAA
TTTTAAGGAA	CATAAATTAT	CTTTTATAGT	TCTGGAGGCT	AGGCAGCCCA	

E7

TGGTGCATTG	AACACATGCA	TGTGCTTGGT	GCCAGCCTGT	CCCAGGTGCT	GAGGGAGCTG
CCTTGGTTTC	CCTAGCAGGG	CTAAGTCTCA	GTGCCACACT	CAGGGAGACA	CTAACGGAGC
ATACCGCTGA	GGCGGCCCCCT	CTTCCTGCAG	GGCCTGTGGG	AGCAGTGCCA	GCTTCTGAAG
AGCACCTCGG	TGTTTGCCCG	CTGCCACCCT	CTGGTGGACC	CCGAGCCTTT	TGTGGCCCTG
CTTGAGAACT	TTTGTTGTTG	GTTGCTGGG	GGGCTGGAGT	GCGCCTGCCC	TGCCCTCCTG
GAGTACGCCC	GGACCTGTGC	CCAGGAGGGA	ATGGTGTCTG	ACGGCTGGAC	CGACCACAGC
GCGTGCAATA	AGTCGGCCCC	CTGCCCCGTC	CTGCCCTGCC	GGGGATGAAC	GGTCTGTCTT
GGGTGGTGTG	CCTTAGGGTG	CTTCGGGGCT	GTGTCACGTA	TGTGCGGCTT	TACCACACCC

E8

GGGGGGCAGC	TGGCTCACTG	GGGGTGAGAA	GGACCTGGAA	GCCTGAAGAC	AGAGGGGAGC
AGTCAGAGTG	GGCACGAGAG	GCTCAGGCTG	TGGCATGGCT	GGTGAGATGA	TGCACCGGTG
GGACCTGCCC	TGGGTAGACC	CCTTTGATGT	TCCTTTTCAG	GCCCAGTGTG	CCCTGCTG
ATGAGGTATA	GGCAGTGTGT	GTCCCCTTGC	GCCAGGACCT	GCCAGAGCCT	GCACATCAAT
GAAATGTGTC	AGGAGCGATG	CGTGGATGGC	TGCAGCTGCC	CTGGTAATGA	ACTTCCCCT
TTATTTACAG	ATCAGAGACC	TTGCCAGCAC	CCTTGGCTTT	GGGAAGTGAG	ATGAAGGCCA
CTTGTGTTTG	CTCCCCGTTG	CTGAAATCAG	AAGCCTTTTA	TTAAGTTTTT	AAATGTCTTT

E9

CCTTCCTTAT	ATTGCATTAT	GTGAAAGATA	AACACCACAG	AACAAGTTCT	TTGAGCTTCC
TGGAAGAAAC	CCAACCATG	TCCCTGGGGA	TTCTATAGTT	GTGGGATGAG	TGACGCAATG
ACAATGTTGA	GGCCTTTGTC	TTGATGCCCT	TGACCCCAAG	GGGACAGCTC	CTGGATGAAG
GCCTCTGCGT	GGAGAGCACC	GAGTGTCCCT	GCCTGCTTTC	GGGAAAGCGC	TACCCCTCCCG
GCACCTCCCT	CTCTCGAGAC	TGCAACACCT	GGTAATGGGG	GCTGCGCAGC	GTGCTCTGGG
AGACCTGCCT	GGGGGACTGG	GGAGGGAAGC	TGGGTCAGCA	GGGGTGGTGG	CAGGTGGAAA

E10

CACATCCCTT	CGTTTGGGGA	CTGTGATAAC	TACCATGAGC	TCTAAATCCA	TTTGCATACC
CTTGTGTTTG	CAGAACCACT	AATGACCTGT	GCTTTTTTCC	TCCAACAGCA	TTTGCCGAAA
CAGCACTG	ATCTGCAGCA	ATGAAGAATG	TCCAGGTAGG	CGACCTGCCG	CTCATTCTCT
TCCTCCTTCC	CTGAATCGGG	GAGGCGTCTC	CTCCTCTCTT	ACCAGACAGG	GAGGGCCACA
ACTCAAGGCA	GCGAGAAGTT	GCCCTCTCTG	CACAGGTTGG	CTGCAACCAC	GTGACCCAG

E11

GGTTTGGGCA	AAGGACGTCC	ATGCAGTTTT	GGGGAAGGGC	ACCCTGCTTG	CATATGCATT
CCACCTTGGC	CACCCAGGG	GAAGTGCCCT	CACCTCCCAT	TCTTCTCCCT	TCCTCTGTGT
CTCTCCAGGG	GAGTGCCTTG	TCACAGGTCA	ATCACACTTC	AAGAGCTTTG	ACAACAGAT
CTTCACTT	AGTGGGATCT	GCCAGTACCT	GCTGGCCCGG	GATTGCCAGG	ACCACTCCTT
CTCCATTGTC	ATTGAGACTG	TCCAGGTGAG	CTTTGCCAGC	CCGGCTGCTG	GTGCGGTGGT
AGCCCAAGGT	GCTGCATAGC	TGCTAGGCTG	AATGGGCAGT	CCCTGGAAGG	CGTGGGGCAT

E12

TTATGAGAAG	GCCAGAGGCT	CTCGGGTTGA	GGCCTTTCTC	TGATTAAGAG	GGTCCTGGGC
TGGGGAGCTG	GATAGGCAGG	GGGTGCAGCA	AAGTCACCCT	GTGTTCCCTC	TTGGCAGTCT
GCTGATGACC	GCGACGCTGT	GTGCACCCGC	TCCGTACCGG	TCCGGCTGCC	TGGCCTGCAC
AACAGCCTTG	TGAAACTGAA	GCATGGGGCA	GGAGTTGCCA	TGGATGGCCA	GGACGTCCAG
CTCCCCCTCC	TGAAAGGTAT	GCTTCGTCCT	GCTCCATCAG	GCCTGGGGCT	GGCACAGCCC
ATCCCTTAGC	ACCCTCCTTC	TCAACCCTGG	CCACTGTCCT	TAAGGACCTG	CATGCCTCTT

E13

CCTCATGGCC	ACCCCATCTG	CCCCTAAGTC	ATTGCTCTTC	AGTGCTACCA	TCCTTTTGAG
ACACCCCAT	TCCTCCCAA	TACATCTGCC	TGCCACCACC	CTGTCTCTC	CCCACCTCTG
CCTGAGTCCT	GTCCTGCTGG	GTTCCAGGTG	ACCTCCGCAT	CCAGCATACA	GTGACGGCCT
CCGTGCGCCT	CAGCTACGGG	GAGGACCTGC	AGATGACTG	GCATGGCCGC	GGGAGGCTGC
TGGTGAAGGT	AGGTGCCCTC	ACGGGGTACT	GGCTCCCTGC	GGCCCGACCC	TTACAAAGTA
CCCCTTGTGC	TCTGGGTAGA	ATGGCTTTGT	GTGGTGGGAG	AAGAATTCTT	TTCTCCTCAT

E14

CAGGGATGGA	GCTGGTAAAT	TTACACTGCA	GTTCACTTTC	CCCCGCCCCC	CCTTTTGAG
CCTAGGGCTT	AGCTGCCTTT	TATTACATTC	ATATTTATGT	TTCACCCGGG	GAACTTTTTT
TTTTTTATTA	AAACCATCTC	ATTAGCTAAA	CAACTATGCC	GCTGCTTTCG	CGGCAGCGGC
TCCAGAGTGG	CCTGGTCTCT	CCTGCAGCTG	TCCCCCGTCT	ATGCCGGGAA	GACCTGCGGC
CTGTGTGGGA	ATTACAATGG	CAACCAGGGC	GACGACTTCC	TTACCCCTC	TGGGCTGGCG
GAGCCCCGGG	TGGAGGACTT	CGGGAACGCC	TGGAAGCTGC	ACGGGGACTG	CCAGGACCTG
CAGAAGCAGC	ACAGCGATCC	CTGCGCCCTC	AACCCGCGCA	TGAGTATGTG	AACCCGGGGG
CAAGGCAGGG	ACCTCGCAGA	ATCTCGAGGT	CTCCACATTA	GTGCAGTGCG	TTCCGGAGGC

E15

CGTTAGATGA	TTGTAAGATA	AATGGACTTG	GGGGTGGGGG	TGGCATTTCG	AGATTTCTCT
GCGTGACCTT	GCAGCCCTCT	ATTAGCAGCA	CTGGGCTATT	TCCAGGGGAG	TGTTGACCGG
GAGGGCGTGG	CCCCACTTCC	TCCCCACCAC	ATCCCAGGCT	CGCTCTCTC	GCCCCACAGC
CAGGTTCTCC	GAGGAGGCGT	GCGCGGTCCT	GACGTCCCCC	ACATTTCAGG	CCTGCCATCG
TGCCGTCAGC	CCGCTGCCCT	ACCTGCGGAA	CTGCCGCTAC	GACGTGTGCT	CCTGCTCGGA
CGGCCGCGAG	TGCCTGTGCG	GCGCCCTGCG	CAGCTATGCC	GCGGCCTGCG	CGGGGAGAGG
CGTGCGCGTC	GCGTGGCGCG	AGCCAGGCCG	CTGTGGTGCG	TGCCCTCCCT	GCCCCGAGCC
CTCCGGGCCG	CCCCCAAAT	CCGTCCACGT	GTGCTTTTTCG	AAGCCCTTTC	TCTGCGTTGT
TTCTGTGGA	AAGAGGGCGT	AGGGCATTGT	GACAAACACG	GGTGGGGAGG	ACGGTTCATT

E16

GATCTTTCTG	CATTACTTCT	TGTAATTGCA	AGCAAGTCTA	CAATTATCTC	AATAAAAATT
TTAATTAAAT	AATTGGGGGT	CACAGCTACA	AGGGGTGGCA	AGTCCTAGAA	CCACAGTCCT
TGCTGTCCAA	CATTCCCGCT	GAGGCCTTAC	TTCTCCTCTC	TCTTCTAGAG	CTGAACTGCC
CGAAAGGCCA	GGTGTACCTG	CAGTGCGGGA	CCCCCTGCAA	CCTGACCTGC	CGCTCTCTCT
CTTACCCGGA	TGAGGAATGC	AATGAGGCCT	GCCTGGAGGG	CTGCTTCTGC	CCCCAGGGC
TCTACATGGA	TGAGAGGGGG	GACTGCGTGC	CCAAGGCCCA	GTGCCCCTGT	TACTATGACG
GTGAGATCTT	CCAGCCAGAA	GACATCTTCT	CAGACCATCA	CACCATGTGG	TAAGTGCAGG
CAGCAGTGTC	AGGGACCTCT	AAAACAGCAG	AGCTGGGGAG	GAAAACGGGA	CCCCAAGTC
CTTTACTTCA	TGGATGGGTA	AACTGGGGCC	CAAGGAAGGA	AGTGACTTGT	CCAGAACTGT
ACAGCAAGTC	AGAAGCGGAG	CTGGGATTTG	GTCCAGGTCT	TCTGACTTCC	TTACTACACA

E17

AAGTCCCATG	GGATTTAGTG	ACGTGGGGAT	CATCCATTGG	TAACGTTAGC	AAGCTGTGCT
TCAGGAGGGG	TTATGGGACT	GGGACCTGGT	TGGAAGGGGC	AGAGAGTGAG	TGGGAGGTGA
AGATGTGGAG	GCAGCGAGTA	TAGACGAGTC	TCGTGAAGCT	C GGCTATGAT	T TTCTTCTCT
G CAGCTACTG	TGAGGATGGC	TTCATGCACT	GTACCATGAG	TGGAGTCCCC	GGAAGCTTGC
TGCCTGACGC	TGTCCTCAGC	AGTCCCCTGT	CTCATCGCAG	TGAGTACTGT	CCCCCTGGAA
GGCCCATTTGA	CTCCATCCTG	CCCAGATTCC	TCACGTGTGG	AATGGCGGGA	GAGAGCTGGG

E18

TCCTGATTCA	ATGTAACCTG	CTATAGCAAG	GTTAAGGGTG	ATGACAAGTA	TGGGACATTA
GATGAGAGCT	GGGGTCAGAA	GAGCTGCCAC	CAATGTAAGC	CAGAGGTCAG	AAGCCCAGGT
GAGAAGATGC	CCTCCCAGTC	CCACACAGGG	ACCCTGGCTC	AGGCAGCTGC	TGGTCCCCGT
GAGTGGGCAA	CTCTGAGTCT	CTTGAATTTA	GTCACAGACT	CTAGGGGACC	AAAGGACAGT
GTGGAAGGTA	GGTCCATTCT	CTCCTTCACT	AATCATCTCT	TTGCTTTTCC	TACCT TCGAG
G CAAZ AGGAG	CCTATCCTGT	CGGCCCCCCA	TGGTCAAGCT	GGTGTGTCCC	GCTGACAACC
TGCGGGCTGA	AGGGCTCGAG	TGTACCAAAA	CGTGCCAGAA	CTATGACCTG	GAGTGCATGA
GCATGGGCTG	TGTCTCTGGC	TGCCTCTGCC	CCCCGGGCAT	GGTGAGTCAC	CAGGCACAGA
GCTGGTGCCT	GCCCTTCAGT	TTTCTTGTAG	GCAGGGATGA	GTGAGGGGGC	TGTGCAGCCA
GAGAGAGGAA	GAAACAAGGG	CTAAACACGG	ATAGCAATGG	AGGTGGTGCT	GGTAATGGGG

E19

TGGGCTTACC	CGTAGGCTCA	AGTCTCAGAC	AACACTTCCT	GGAGCGGGCT	GGAGGAGGGC
TTTAGATCAG	TCACTGTGGC	CCTGAGGACT	TTTGATTCT	TTTCTCTTAG	G TCCGGCATG
AGAACAGATG	TGTG GGGCTG	GAAAGGTGTC	CCTG CTTCCA	TCAGGGCAAG	GAGTATGCCC
CTGGAGAAAC	AGTGAAGATT	GGCTGCAACA	CTTG GTGAGG	CTCAGTGAGG	GGCTGCGCCG
GGGACCCAGG	CCCTGCGGGT	GGAGTGAGGG	TGCACGCGGC	CACAGGACCT	TCCGCACTTG
CCTCCAGCCC	CCTGTGGGAA	CCTGGTTACT	CTCTGCCTTT	CCTCCGTGCC	TGTGTTTCTC

E20

TTGGGCTTCC	TGGAGGGTTG	GGAGGCACCC	AGAGTATTCT	GTGTTCCCTC	ATTGCCTCCA
TGGCCCAGAT	GGGCCCCTCA	AACCCAAGGT	GCCCAACTTG	TCATCTCTGC	CATGACTGCT
CCTAG TGTCT	GTCAGGACCG	GAAGTGGAAC	TGCA AGACC	ATGT GTGTGA	TGCCACGTGC
TCCACGATCG	GCATGGCCCA	CTACCTCACC	TTGACGCGGC	TCAAATACCT	GTTCCCCGGG
GAGTGCCAGT	ACGTTCTGGT	GCAGGTGAGA	GGTGGGGAGA	TGGGGAGAGG	GTGCTGTTTC
TTTCTAGGAG	GGGTGGGAGG	TGTGGCCTCA	GGTTGGGTTC	TGTGGATCTG	TCTGCAGAAA
CAACTCTGGG	GTCTGGTTTC	TACTGGAGTA	CTTCCAGTC	CTTCACAGAA	GTGCCTGAAG

E21

CTGGGCATAT	CCCCGTCCC	CCAGACACAC	ACAGAGGCAC	ATATGCGCAG	CCATGGACGT
GGCAAGATCC	TGTGACACGT	ACTCAAAGGC	CTGTGATGAA	GAGATGCCAA	TCTTCTGGTC
TGGTGAGAGC	CAGTGGGGAT	AATGGTCTTC	TCCTGGCACT	CCTCTTTCCC	CAG GATTACT
GCGGCAGTAA	CCCTGGGACC	TTTCGGATCC	TA TGGGAA	TA TGGGATGC	AGCCACCCCT
CAGTGAAATG	CAAGAAACGG	GTCACCATCC	TGGTGGAGGG	AGGAGAGATT	GAGCTGTTTG
ACGGGGAGGT	AAGTGCAGCC	TCATCTCCAC	CCTCATGTCC	CGCTTTGTGC	TTCTGCCACT
TAATAGGAAC	ATTTTCCAAG	CATTCAATTA	GAGCTCGTGT	GAATGGAATA	ACGCACAGCC

E22

GACTAATTGT	CAATAAAATA	CCTGCTCATT	GGTTTCATAG	TACTTTAATT	TCATAATCAT
GATTTTGCTG	CTACCTCTGT	TACCGTCTCT	TGGTCATGGA	TGCCTGGAGA	GTGGTGGTGG
TGAGATGGTC	ACAGACATGT	CCTGGCGTGG	GGCTGGCCCT	GCAGGGGTGC	AGTGGCAGGT
GGGGTCCTGG	AGGGGTGGCA	GTGCCTGCAC	TCGTGGGCAC	TGAAGACAGA	TGGGCAGGTG
TAGAGTGGAG	GGAGGATCTG	GCTGTCGAGC	CTGCCCTTCA	TCCTCCTGGA	TTTCTTGCTT
TGTCTTCCTC	CAGGTGAATG	TGAAGAGGCC	CATGAAGGAT	GAGACTCACT	TTGAGGTGGT
GGAGTCTGGC	CGGTACATCA	TTCTGCTGCT	GGGCAAAGCC	CTCTCCGTGG	TCTGGGACCG
CCACCTGAGC	ATCTCCGTGG	TCCTGAAGCA	GACATACCAG	GTCAGTGGCT	TTCTTGCTTC
ATCTTGTTGG	GGACTTGGCC	TTTGAGTGT	TTTCTGCTCC	CTGATCGTAG	GTCTCTAAGG
ACTTGCTTTA	TGAATCCAGG	TGCTCCTGTG	TTGGGTGCAT	ATATATTTAG	GATAGTTAGC

E23/24

TCTCAGGGCC	AAGGCTGCCT	GATTCCCACC	TCTGCCCTTG	GCTGACTATG	TGACATGGGC
ATGTTGCCCTC	TCTGTTTCCA	TAGCTTTAAA	TAAAATGGGG	CCAGCAAGGA	AGCTCAGGAA
TGGGTCTTGG	CAATGGCAAG	GCTTTGCTGC	TCACCTCGGG	CCTCCTCTGA	GTCTCTGTCC
CGCTCCTCCT	CCTCTTCCTC	GAATGCCCTC	TGCCTCCATT	GCCGCCAGGA	ATGTTCCCTC
TTCCCTGAG	CCGGAGAGCA	TGCTCCTGGG	CTTGACGGTG	CTCATCCCTC	AACTTGTCTC
TCAAGGAGAA	AGTGTGTGGC	CTGTGTGGGA	ATTTTGATGG	CATCCAGAAC	AATGACCTCA
CCAGCAGCAA	CCTCCAAGTG	GAGGAAGACC	CTGTGGACTT	TGGGAACTCC	TGGAAAGTGA
CTTGGCACTC	TGCTGACACC	AGAAAAGTAC	GTCTGGGTCT	CTGTGTGGAC	AGAGCCCTAG
AGCTTGCTTC	CTGGAATGTC	CCTCTGTCCC	CATTGTTCATG	GGGGCTGGAA	GGGGGGTTGT
GGGTGGTATG	ACCTCCAGGT	GGCTGCAGGG	TGGGAAGGAG	GGTCTCTTGG	ATCCTTCTGG
GCTGAATAAC	CCCAGTTTGA	CCAGCTGACG	GCTGGCCTAT	CTCTTGCTTG	GTTCCAGGT
GCCTCTGGAC	TCATCCCCTG	CCACCTGCCA	TAACAACATC	ATGAAGCAGA	CGATGGTGGA
TTCCTCCTGT	AGAATCCTTA	CCAGTGACGT	CTTCCAGGAC	TGCAACAAGC	TGGTGAGGAC
CTTGAGGGTA	GTGGGAAGCA	GACGGTCCCA	AGGCTTGGCC	TGGTGGTATG	GACACAGAGT
GTGACCTTCT	AACGTGGACA	CTACCCTCGT	GTCTTGACAT	GATCTGCACC	AAGACACCAC

E25

GGATTTGGAA	TCAGAGTCTA	GGCCCTATTG	TCCTGCATAC	CTGCCCCATA	GCACTGCAGT
GAGGCAGCTG	CAGGGCCTGA	TGTATTTCCC	CATTCTCTTT	GCTGAATTGA	GGCAAAGAAA
GACAGTGACC	AGCACATATG	TGTGTTTGTG	TTTTTGTAAG	AGCACCCACA	TGCTCATGAG
GCTAAGAGTG	GGTTGTGAGG	ACAGATGGGT	GGCTGAGCAG	GGAGGTAGGC	AGAGGGACAG
AGGGAATGTT	CTTCTGGAAA	ATCCTCAGGC	TCATTGTGTT	CTGCAGAAGG	CCAGCAGCAC
TGCATTATTC	AACTCTTCTT	GCTGGAATGC	AGATTAGAAA	CTAAGAATCT	TGCCTTCCCA
CTCATTCCCT	CTTTGAGACC	ATTGAGCTAC	ATTTCTCCTT	CTACCTGGAC	CCCCTTATCC
TTAAATTGAC	CATCAGAACA	TTTGCACCCA	GACTAAGAGC	CAGAGTTTCT	GACACCTGGC
CATAGGCCTG	GGCCACCTGA	GGCTGCCTTT	GCAGGTGGAC	CTCGAGCCAT	CTCTGGATGT
TGCATTTTAC	GACACCTGCT	CCTGTGAGTC	CATTGGGGAC	TGCGCCTGCT	TCTGCGACAC
CATTGCTGCC	TATGCCACAG	TGTGTGCCCA	GCATGGCAAG	GTGGTGACCT	GGAGGACGGC
CACATTGTGC	CGTGAGTACT	GACGCCCTCA	TGTTCTCAGA	TGCCCTCCCT	TCTTCCCATG
TGTCTATGCT	TGAAGACCTT	GTGAGTGCAG	GGGGATATCT	TCATGGGCGA	GAGAAGACAG
AGTGTTAGTA	GGGACTGGAT	GGCCAAGGGC	TAAGGAGGGT	TAAGACATTG	GCTGTGTAAG
AAGTTTATAT	TACGGGTGAG	GTGGGACATG	GATTCAAGGC	ATGAACATGT	GGAGACTTTT

E26 – Ligation site is 367 nt before exon 26
(Probe sequence: TAGAATCTTG-CTTCTTTGGA)

CTTCAACAAT	CCAGAAGTCT	CTAATATTGG	TGACGCCCAT	AGTCCCCTAG	TTCCCCAACA
TTATCTCCAG	ATGGCGCAGG	CCATCACCAC	ATGGGTCTGC	AGTCCTGGAG	GCTTTGCCTG
TTGTGGCCAC	AGCCTTGTCT	CCTGTCTACA	CAGCCCAGAG	CTGCGAGGAG	AGGAATCTCC
GGGAGAACGG	GTATGAGTGT	GAGTGGCGCT	ATAACAGCTG	TGCACCTGCC	TGTCAAGTCA
CGTGTCAGCA	CCCTGAGCCA	CTGGCCTGCC	CTGTGCAGTG	TGTGGAGGGC	TGCCATGCCC
ACTGCCCTCC	AGGTGAGGCC	TCTATCCCTG	GGGGTCAGGC	TGGTGGGATG	GGATAGGGAT
GGATGGAAAG	GTGCTTCTAG	GTCTTGCTTC	ATCTCAGCCT	CCACCTGCCA	CGTCTATCT
CTGACCTGCA	AGGCTGCTGC	AGGTTCCGTG	GTTTCTTTCA	TCAGAGTCAG	GACAGTCGTG
ATTTTCTCA	AGTCGAGCTC	CTCCAAAATG	CTTTTCTGTG	CCTATTTATG	GGATTCTCAC

E27

CCCGGAGATA	AGATTTTCATT	TGTCCTGAGG	CTTGGGGAGG	TGAGGTAGGG	TGAAGGAATC
CCCGCTCCCA	GTTTTGCAGA	GGGATCAATC	AAGGCACAAG	CAGGAGAGAT	GCTCCTTGAG
TGATGGGGTG	ACCCCTGGGA	GTGCAGGCAG	GAGGAGTTGG	CTTCTAGGGC	AGGAGGAGGA
GTTGGCTCCT	CCCTTTTAGT	TAAAAATGAG	GCTTCCTCGT	GGGAAAGGGG	AGCGTTTTTG
TTCCTAATGA	GAGCTTTCTT	TTGCAGGGAA	AATCCTGGAT	GAGCTTTTGC	AGACCTGCGT
TGACCCTGAA	GACTGTCCAG	TGTGTGAGGT	GGCTGGCCGG	CGTTTTGCCT	CAGGAAAGAA
AGTCACCTTG	AATCCAGTG	ACCCTGAGCA	CTGCCAGATT	TGGTAAAAACA	GATTCCTGGG
TTGTTTGAAG	TGATGAATCT	TATTGCTTCT	CCAGGGGCCA	GCTTAGAGAC	TAGGTTTTTG
GTAAAAAGTC	CTAGCCACAT	GAGTTGAGAG	ACTGAGTTTA	TTTTATTTGT	TTATTTATTT

E28

TAGGATGGGC	TAAAGCTTTG	GTTTCTAGAA	TCAAAGGAAG	TCGGCTATGT	GTGTGTTTTG
AAGGTGGGGG	GCATGCTATT	TGGGGACAGA	TGTTAAACAA	TGACATCTCA	CTTGGATGTG
GAATGGTCCA	TGGGATCTCA	AGTTCAGGTG	GAACAGAGGA	GATTCTGTGG	GAATATGGAA
GTCATTGTAC	ACTGTAGGGC	TCAGAAGTGT	CCACAGGTTC	TTCCTGAACC	ATTTTAATTT
CTTCGCTCTT	TTCTGCAGCC	ACTGTGATGT	TGTCAACCTC	ACCTGTGAAG	CCTGCCAGGA
GCCGGGAGGC	CTGGTGGTGC	CTCCACACAGA	TGCCCCGGTG	AGCCCCACCA	CTCTGTATGT
GGAGGACATC	TCGGAACCGC	CGTTGCACGA	TTTCTACTGC	AGCAGGCTAC	TGGACCTGGT
CTTCCTGCTG	GATGGCTCCT	CCAGGCTGTC	CGAGGCTGAG	TTTGAAGTGC	TGAAGGCCTT
TGTGGTGGAC	ATGATGGAGC	GGCTGCGCAT	CTCCCAGAAG	TGGGTCCGCG	TGGCCGTGGT
GGAGTACCAC	GACGGCTCCC	ACGCCTACAT	CGGGCTCAAG	GACCGGAAGC	GACCGTCAGA
GCTGCGGCGC	ATTGCCAGCC	AGGTGAAGTA	TGCGGGCAGC	CAGGTGGCCT	CCACCAGCGA
GGTCTTGAAA	TACACACTGT	TCCAAATCTT	CAGCAAGATC	GACCGCCCTG	AAGCCTCCCC
CATCACCCTG	CTCCTGATGG	CCAGCCAGGA	GCCCCAACGG	ATGTCCCAGG	ACTTTGTCCG
CTACGTCCAG	GGCCTGAAGA	AGAAGAAGGT	CATTGTGATC	CCGGTGGGCA	TTGGGCCCCA
TGCCAACCTC	AAGCAGATCC	GCCTCATCGA	GAAGCAGGCC	CCTGAGAAC	AGGCCTTCGT
GCTGAGCAGT	GTGGATGAGC	TGGAGCAACA	AGGCGAAG	ATCGTTAGCT	ACCTCTGTGA
CCTTGCCCCCT	GAAGCCCCCT	CTCCTACTCT	GCCCCCGAC	ATGGCACAAG	TCCTGTGGG
CCCGGGGCTC	TTGGGGGTTT	CGACCCTGGG	GCCCAAGAGG	AACTCCATGG	TTCTGGATGT
GGCGTTTCGT	CTGGAAGGAT	CGGACAAAAT	TGGTGAAGCC	GACTTCAACA	GGAGCAAGGA
GTTTCATGGAG	GAGGTGATTC	AGCGGATGGA	TGTGGGCCAG	GACAGCATCC	ACGTACCGGT
GCTGCAGTAC	TCCTACATGG	TGACTGTGGA	GTACCCCTTC	AGCGAGGCAC	AGTCCAAAGG
GGACATCCTG	CAGCGGGTGC	GAGAGATCCG	CTACCAGGGC	GGCAACAGGA	CCAACACTGG
GCTGGCCCTG	CGGTACCTCT	CTGACCACAG	CTTCTTGGTC	AGCCAGGGTG	ACCGGGAGCA
GGCGCCCAAC	CTGGTCTACA	TGGTCACCGG	AAATCCTGCC	TCTGATGAGA	TCAAGAGGCT
GCCTGGAGAC	ATCCAGGTGG	TGCCCATTGG	AGTGGGCCCT	AATGCCAACG	TGCAGGAGCT
GGAGAGGATT	GGCTGGCCCA	ATGCCCTAT	CCTCATCCAG	GACTTTGAGA	CGCTCCCCCG

AGAGGCTCCT	GACCTGGTGC	TGCAGAGGTG	CTGCTCCGGA	GAGGGGCTGC	AGATCCCCAC
CCTCTCCCCT	GCACCTGGTA	TGCTGGCACC	TTGTGTGCAG	GTGGGAGGGC	TGGGCGAGGG
CTGGCATGGC	CTTGGTGCTA	CATGCATCTG	CCAAGATACG	ACTCGGGTTC	TAATCCTGGC
TTCCCTGGTC	TGTGTGGCCT	TGGTTGAAAC	TTGCCTTCAA	AGGGCCTGTG	TTTCTCACC
TCCCTGGCAG	GGAGACAAAC	TGTGATCCTT	TTTCGGGGGA	AGATCTTTTC	TTATTGTGTG

E29/30

ACCCCCCGCC	CCCCTTGGTG	GTGCCACGTG	GTCAGCACGC	CCTGCAGATC	CTATTGGATG
TCAGGTTGTA	GGCCTGGTGG	CCATTGTCCC	TGCTGGCACC	TGTGTGCTCA	CCTTCCTGGT
TGTCTTTGCA	GACTGCAGCC	AGCCCCGGA	CGTGATCCTT	CTCCTGGATG	GCTCCTCCAG
TTTCC	AGGG	TTTTT	ATGAA	ATGAA	GAGTTTCGCC
CAATATAGGT	GGGTGAGCGA	GGCACCTGAA	GCAGCAGGTG	ACGAAGAGGC	TCTTTTTGTG
GCTCTACTTG	ATTCAAATA	ATCCGCATTT	TCTCGTTCCG	TTTAGGGCCT	CGTCTCACTC
AGGT	AGGG	CACTTA	CGA	GCATCA	CCACCATTGA
CGGAGAAAGC	CCATTTGCTG	AGCCTTGTTG	ACGTATGCA	GCGGGAGGGA	GGCCCCAGCC
AAATCGGTAA	CGTTGGTGCC	ACAGGCTGGA	TGCAGAAGCT	GCATTCTGGT	TCTTATTTTT
GGCATAAGTG	ACTGTGTGAC	CTCGGCCAGT	CACTTTGCTC	CTTGGCCTTA	GTTTCTTCTC
CTGGAAAGTG	AGGGGCTAGA	TGCTCTTCCA	CGTCTCTCCA	GATCTCAACT	GGGTGTTCTT

E31

ACTGTGTGAC	CTCGGCCAGT	CACTTTGCTC	CTTGGCCTTA	GTTTCTTCTC	CTGGAAAGTG
AGGGGCTAGA	TGCTCTTCCA	CGTCTCTCCA	GATCTCAACT	GGGTGTTTCT	TGGAGTTTCT
GAATCATTCA	GCTTTTAAGT	GACTTAAGGA	TCCACCGTTA	AGACAGGGTG	TCGAGCCGCA
GTCAGTACTG	ACTTGGCGTG	ATCTGTTCTC	CATCCTCAGG	GGATGCCTTG	GGCTTTGCTG
TGCGATACTT	GACTTCAGAA	ATGCATGGTG	CCAGGCC	AGGCTCAAG	GGGGTGGTCA
TCCTGGTCAC	GGACGTCTCT	GTGGATTGAG	TGGATGCAGC	AGCTGATGCC	GCCAGGTCCA
ACAGTAAGAA	TCTGGTGTAC	AGTCCTCAAT	TCAGGAGAGC	GAGACTTCAG	AAATGATAGA
GGATTAAAAG	TGGGCTGGGG	TTACTTTTGG	ATGTTGTGGC	TGATATTAAA	TATGCTCAAT
ACCAACCTCT	GTCTGGTTT	CAAACAGCAG	GAAATTATCC	AACCACGATG	TGATTAGGAA
ACCAAACAAT	AAAAAGTGTG	TATTGATCCT	ATATCTTTAT	AACTAATTAT	AATCTTTAAT
TTCTCCACAT	TTTCCCTTTG	TGTCCAGGAG	TGATCATACC	TCTATGCCCT	

E32

TAATATCTGT	TGACTGAACA	TCTTCCTCAT	AGGGCTGATG	TATGTGACCA	GCCTGGAAAA
CATGAGGCTG	TATTCAGATG	CTGGATATAA	CGTCAGGCCA	GTCCATTTTG	AGCCTTCTTG
CCCACAGATC	CTTTCTTGT	TCTTTGCTAA	CTCTAGGAGT	GACAGTGTTT	CCTATTGGAA
TTGGAGATCG	CTACGATGCA	GCCCAGCTAC	GGATCTTGGC	AGGCCAGCA	GGCGACTCCA
ACGTGGTGAA	GCTCCAGCGA	ATCGAAGACC	TCCCTACCAT	GGTCACCTTG	GGCAATTCTT
TCCTCCACAA	ACTGTGCTCT	GGTGAGTCTT	ATAATACCTT	TCTTACTTCC	CTCAAAAAAA
AGTTCCAAAA	TAAACCCGCC	AAAGAAACAA	AAATAGTATT	CAAGAGACCC	CAGGGTGCCC
ATGCATAAGA	TTTGGCCTTG	AATGAAGATT	ATTATATAGA	CCAGGACCTT	TAAGTTTCTG

E33

TGGCATCATA	ATTATTAATA	ACTGCAGAGC	ATGTCTGAAG	AACCACAGGA	TGACCACCTC
AGCCTCATGT	CCCTATGTCT	CCACTGTTAA	CCTTGTTTCA	ATTCTTTTCA	GAGTTGAGTT
GACTTCAAAA	ACTAGACCAG	GTTGCTTAAG	CAGACATTGT	GAATGGTTCA	GAATTTCTGG
GTGAAAGATG	GGAAC TAAG	TCTTATTTGT	GTCTGTTGCA	GGATTTGTTA	GGATTTGCAT
GGATGAGGAT	GGGAATGAGA	AGAGGGTAAG	TTCCTTTCTG	TTGACTTTGA	AAGAAAGGTT
AGAGATGTGT	TTGGGGCTCT	TGTTCCCACT	GGTTAATTTT	TCCTCCTTTG	GTCTTAGTCC
AGTGCTTCCT	TTTACTATTA	TCTTGTTTTT	GCGGGTCCAT	CTGTACATCT	TGTGTTTTGC
TTCTGTCTC	ATGTACAGGG	GGCCTCCTTG	CTGTGTAGGC	CTGTGTTCAA	TTCTAGGGGT
CAGTTGTCTG	GCAGATGGGC	TTAGAGTTGG	AGTACCTCAT	CTTATTCCCT	GCCTGAATCT

E34

GTCTTAGTCC	AGTGCTTCCT	TTTACTATTA	TCTTGTTTTT	GCGGGTCCAT	CTGTACATCT
TGTGTTTTGC	TTCTGTCTC	ATGTACAGGG	GGCCTCCTTG	CTGTGTAGGC	CTGTGTTCAA
TTCTAGGGGT	CAGTTGTCTG	GCAGATGGGC	TTAGAGTTGG	AGTACCTCAT	CTTATTCCCT
GCCTGAATCT	GCTGTTTTCT	TCTGCAGCCC	GGGGACGTCT	GGACCTTGCC	AGACCACTGC
CACACCGTGA	CTTGCCAGCC	AGATGGCCAG	ACCTTGCTGA	AGAGTCATCG	GGTCAACTGT
GACCGGGGGC	TGAGGCCTTC	GTGCCCTAAC	AGCCAGTCCC	CTGTAAACT	GCAGCAAGCC
TGTGGCTGCC	GCTGGACCTG	CCCCTGTGAG	TCCTTTGCTT	CTCCAGCCAG	GGCAGCGTCA
AAGGGGCAGT	GCTTTTAGCT	TGGCTGTGCA	GAAAAGTAGA	GCAGGCACCC	ACCAGCCAG
AAGTACCCTT	TCCCTCATCA	CCACATGCAC	AGTGCTACCT	TCACTCACCT	TCCTTTCCCTC
CTGTGCTCTT	TGGACATGCA	TGCAGCCAGT	CTCAGGGATC	ACTGCCCTCT	TTCTCTGTCT
TTGGAAGGCA	CTTCCCCAGA	TTATGCATAA	CTGGAAGGAA	GAATTGCTTT	TCTGAGGTCA

E35

AGTTTCCTAA	TCTGTAAAAA	AAGGAGGAAA	TTGTTCTATA	TGACCTCAAA	GGGCCTGTTC
CGTTCTCTAC	TGTATTTATC	TGTGTGCAAC	TTGGTCACAC	CTGCCTGTCT	GCATGTAGTA
GGCATGGGGG	TTTGGATAAC	GTCGCATCCA	TCCTCTGCTT	CTCTCTGTCC	AGGTGTGTGC
ACAGGCAGCT	CCACTCGGCA	CATCGTGACC	TTTGATGGGC	AGAATTTCAA	GCTGACTGGC
AGCTGTTCTT	ATGTCCTATT	TCAAAACAAG	GAGCAGGACC	TGGAGGTGAT	TCTCCATAAT
GGTGCCTGCA	GCCCTGGAGC	AAGGCAGGGC	TGCATGAAAT	CCATCGAGGT	GAAGCACAGT
GCCCTCTCCG	TCGAGCTGCA	CAGTGACATG	GAGGTGAGAA	GTACTTTCTG	TGGATCCGTG
GTAAGGCAAT	AGAATGTCAG	GAAAACCACC	TGGACCTGGT	GGCAGTTGCT	TTTAGTTGAT
GCTCTTGTTA	GGAGCTCTGC	CTTCTGCTTA	AGTGAGGAG	AGGAGTACCA	CTTTCTTAGA
GGGGTTTATT	GCCATCCCTT	TGTCTTGCG	TGATTTTCATG	TTGTTCCGGG	CTCAGATTTG

E36

GCATGGCCAC	CCGTACAGTA	TTACAACAGC	CTGTTTTCTG	GTCAAGCTCT	GTAAAGTCAA
GCTCTGTTTT	CTAGTCAAGC	TCTTCTTTAT	CTTAGAAGAG	CTGCCGACAA	ATATCAAGGT
TTGTGGCTGA	TGTTGCAGTT	GAGTGTGATG	AATGTGCAGG	AACTCTCGGT	AACTTCCCTA
AAAACCTAGG	ATTCTTCATT	GCTAGGACTA	CGGATGAGCT	CTTTCTTCTT	TGTGCAGGTG
ACGGTGAATG	GGAGACTGGT	TCTGTTTCC	TACGTGGGTG	GGAACATGGA	AGTCAACGTT
TATGGTGCCA	TCATGCATGA	GGTCAGATTG	AATCACCTTG	GTCACATCTT	CACATTCACT
CCACAAAACA	ATGAGTTCCA	ACTGCAGCTC	AGCCCCAAGA	CTTTTGCTTC	AAAGACGTAT
GGTCTGTGTG	GTAAGAACAT	TTTCTCAACT	CCTCTTCTCC	CCCTGCTATA	CATTTATAAA
CCTTACTTGC	TCTACTCTGA	GGCTCTTGGA	TGCTTATATT	TCAGGGTCTA	GTAGCGAGGG
TCAGATTCTG	GTGAGGATCA	AGAATGGCCT	GTCTCTGGCA	TCAATGTTTT	TGTACCCAGG

E37

GGCTCTTGGA	TGCTTATATT	TCAGGGTCTA	GTAGCGAGGG	TCAGATTCTG	GTGAGGATCA
AGAATGGCCT	GTCTCTGGCA	TCAATGTTTT	TGTACCCAGG	GCCACTCAGT	TTATCTTTTT
TTTGTTTGTT	TGTTTCTCTA	GGGATCTGTG	ATGAGAACTG	AGTCAATCAC	TTTCATGCTGA
GGGATGGCAC	AGTCACCACA	GA CTGGAAAA	CACTTGTTCA	GGAATGGACT	GTGCAGCGGC
CAGGGCAGAC	GTGCCAGCCC	ATCCTGGAGG	AGCAGTGTCT	TGTCCCCGAC	AGCTCCCACT
GCCAGGTCCT	CCTCTTACCA	CTGTTTGCTG	AATGCCACAA	GGTCCTGGCT	CCAGCCACAT
TCTATGCCAT	CTGCCAGCAG	GACAGTTGCC	ACCAGGAGCA	AGTGTGTGAG	GTGATCGCCT
CTTATGCCCCA	CCTCTGTCTG	ACCAACGGGG	TCTGCGTTGA	CTGGAGGACA	CCTGATTTCT
GTGGTGAGTC	TCCAAGTTAC	CTCTGAAAAT	CCTGGAGACC	AGCTAACTGG	GCTTGCTCAG
CCTCTCTGTG	CCCCAGATTG	TTTATCTGGC	TAATAAGGAT	GAAAATCCCT	GTATTTTCGTA
TGACCTGGAA	TTCCAGTTAG	GATCAAATAA	GATGGAAGAT	ATATATATAC	

E38

TACACTTAGC	TGCAAGAAAG	GTTGGGAAAT	ATAGTCCTCA	TTCTGGTGGT	CATATGCCCCA
GCCTAAAGTT	CTGTTTCTAT	GGAAGGTGGG	GAGGATGAAG	ATTGGTGGAA	AATAGCCGTC
TCTGCCCTGG	CAAGTTTGTC	TGATGATTAA	CCATGTTGAA	TCAGCTGTGC	CCATTTCACT
CTGGCTGGTG	TGGGCCTTTG	CAAGTGACCT	CCTTCTCTGT	CTACAGCTAT	GTCATGCCCCA
CCATCTCTGG	TCTACAACCA	CTGTGAGCAT	GGCTGTCCCC	GGCACTGTGA	TGGCAACGTG
AGCTCCTGTG	GGGACCATCC	CTCCGAAGGC	TGTTTCTGCC	CTCCAGCTAA	AGTCATGTTG
TAAGGCAGCT	GTGTCCCTGA	AGAGGCCTGC	ACTCAGTGCA	TTGGTGAGGA	TGGAGTCCAG
CACCAGGTAG	GAGCCTGGGC	CTTTCACTTC	CCATGGGGCT	GCTTCCCTCT	TGGGTCAAAA
GGGAGATATG	TTCACCAGTG	ATTGGCCTCT	TCCAAGTGTG	GCAGCAGGGT	TGACTCTTAC
GGGGATCATT	AGGGATAGGG	AGGCAGTATT	GTGTAATGGC	CTTGCAATTCA	GATAGCTCTT

E39/40

CTCCTTTTCA	CCTGCTTTTG	TTTGCTGTG	AGGCTCTGTC	AGAGGGCTGG	GTCAGGGACT
GGCCCTTGCC	GGAGGCAGGG	AGAGCTGACC	TCTGTGAATT	CTGGGCTTCG	TACCTAGAAT
GTCTGTGCC	CTTTCTGAAC	CTTGCTTTGC	CCTCAGTTCC	TGGAAGCCTG	GGTCCCGGAC
CACCAGCCCT	GTCAGATCTG	CACATGCCTC	AGCGGGCGGA	AGGTCAGCTG	CACATGCGAC
CCCTGCCCCA	CGGCCAAAGG	TGAGAGTCCT	CCCCTCCCTG	GTGCCTTCAT	GGAGGAACAA
GGGCCCTGC	AAGGCCCCCC	AGCCACCCAT	CTTCACCTCT	GGCAGAGCAG	ACTCAAACAC
TGGCACCTAG	AGTCCTAGAG	TGGGTGGGCT	TCCTTGCCCA	GCCTGCATTT	CCCATCACTG
GGCCTGGGAG	CCCCATTCTG	CACCTGGGGT	CGACATTCTC	AGATTAACCC	TCGCCTCTGG
TCCCCAGCAA	CGGTGAGACT	TAAGAGTCCC	CTGGAGGGTA	AATGTGAGGG	TGTCAACAGG
AACATGGGGA	CACTCATCTG	TCAGAGGTCC	CGTGGCCTGG	ATCCTTGTGG	GATGACCGTA
CAGAACTCCT	ACTAGTTTTT	AGTGAGCAAG	AACATTTCAA	ATCCCTCTGA	GGCTGTCCCCA
CCACTAATTT	CTCTGACTTT	TGTGGCCGTT	CCTCTCCTCT	AGCTCCACAG	TGTGGCCTGT
GTGAAGTAGC	CCGCCTCTGG	CAGAATGCAG	ACCAGTGCTG	CCCCGAGTAT	GAGTGTGGTA
TGTGTCCAC	CAGGGGGATG	TCTCCAGGGC	CCAACCCTAG	CCCCAGGGGG	CACCACGTTG
AAGGTGCTGA	AAGGTGTCTC	TGTTCTCAGG	CACAGGGTGT	GTGAAAGGAG	GTGGGTAAAGG
ACCGATTGGA	TACTCCAAAA	AAGTGGAAG	GGTTACCTCT	GGAGAATAGG	ATTTGCTTCC

E41

TATCATTCTA	TCTATCTATC	TATCTATCTA	TCTATCTATC	TATCTATCTA	TCCTAACCCA
TGTAATCTCT	GTCTCCATCA	TCATCACTTA	CCTAAAACAG	TAGAAGTCTG	CATGAATAGG
AATGTAGCAT	CCCACTCACA	GGTAATAAAA	GAGTAACCTT	TCTGAACTCT	GCATGGACGT
CTCTCTTTCT	GGCCCTCAGT	GTGTGACCCA	GTGAGCTGTG	ACCTGCCCCC	AGTGCC TCA
TCTGAACGTC	GCCTCCAGCC	CACACTGACC	AACCCTGGCG	AGTGCAGACC	CAACTTCACC
TGCGGTAAGG	CCTCTGTGGA	TGAGGAGGGG	TGGTGTGGCC	TCTCTCTGCT	GGTGTGAGGG
AGGCCATCCT	CCTCAGGGAC	CTCTTCCAAG	ATCACGTCAT	TTCCTGTTTT	CTACCTAGCT
GAATCTGGGT	TGGGAGTACA	TCTGGAACAG	AGGGGTTAGG	GTCACACCTG	CACGGAATCC

E42

CTTCACATTT	CCCTTTTTGA	TTAATGTTCA	TATTTTTCTG	CAGATGGACT	GCTTTTGGTA
ACATTGAATA	ACTCCCAGCC	CGTGAGCTTG	GCCCTCACAC	ATTTCTGACT	TAATCTTCTG
AGTCTAAAGC	TCCCTGGCAC	CCTATAGCAT	AGCTGAATAC	TTACGAGCCC	TGGCTGGGCG
CAGTGCTCAG	TGTGGCCTTG	TCCTATCCTC	AGCCTGCAGG	AAGGAGGAGT	GCAAAAGAGT
GTCCCCACCC	TCCTGCCCCC	CGCACCGTTT	GCCACCCCTT	CGGAAGACCC	AGTGCTG TCA
TGACTA TAC	TGTGCTTGCA	ACTGTGTCAA	CTCCACAGTG	AGCTGTCCCC	TTGGGTACTT
GGCCTCAACC	GCCACCAATG	ACTGTGGCTG	TACCACAACC	ACCTGCCTTC	CCGACAAGGT
AAGGACTGCT	TGGCTATTAA	CTATCAGTTA	ATAGTTTACT	CATTTATTTA	TTGCTGTCTAG
TTTATCCTTC	TATCCACCCA	TCCATTTCATC	CATCCACCTA	CCCATCCAAT	ATTTGCTAAG
CAACGTGTGC	TCCTCATGGA	AGATTTGCAT	CCTACCCAGC	ATTCCCTTCT	TGCCCCAAACC

E43

GCTGGTGGGC	AGGGAAAGGA	GGAGGACGAA	ATCTGCTTTA	TAAGCAGCCC	TGCATATGGG
CGGAGACCAC	ATTTGGTTTT	ATTTATTGCC	ACATTCTCAG	CACCTAGAGC	GGCTTCTGTG
TAGTAGGTGC	TAAC TCACCG	CTTGTTTCAA	TGAAACAAAT	GAGTTCAC TC	ACGAAAAC TT
ATGTCTACAG	GTGTGTGTCC	ACCGAA TCA	TATCTA CCT	GTGGGC CAGT	TCTGGGAGGA
GGGCTGCGAT	GTGTGCACCT	GCACCGACAT	GGAGGATGCC	GTGATGGGCC	TCCGCGTGCC
CCAGTGCTCC	CAGAAGCCCT	GTGAGGACAG	CTGTGCGTCG	GTGAGTGGGG	CAGGGGCTGG
GCATGCCTGC	AGCTATCAGA	GCGGGAAAGT	AGAGGAGGGC	ATCTTAGGAA	GGGTAAGAAA

E44

GCTGGTTGAC	AACAAAGATG	AAAATGCCCA	GACCAGTGAT	CACTTGGCAC	AACCCAGGG
CTCAGTACGC	AGGAGGCGTA	GGTAAGAGCC	CCTGTGTCTT	TGCTGCTGGC	CTGTCCTTAC
TCTGTTTTTT	CTGCTTTTCC	AGGGCTTAC	TTACGTTCTG	C TGAAGCC	TGTGCTGTGG
AGGTGCCTG	CCATCTGCCT	GTGAGGTGGT	GACTGGCTCA	CCGCGGGGGG	ACTCCAGTC
TTCCTGGAAG	AGTGTAGGTC	CAGGCCCCCG	GGACGGGGAG	GAGGGCAGAT	TGGGGCCACT
CCAGGGACCA	GCGTTGACCT	TGGTTTCATC	TCATTCCCTG	CCCTTTGCTT	TCCCCACTGG
GCATTTTACC	CAGCTGTTGG	TGTTATCCAG	ATGCCAGCTG		

E45

CCACTCATCC	CCTCCGTGGG	CCCTACCCTG	TGGTGGGACT	TACATGTTAA	GCCAGGCTTC
ACGTCTAGAA	ACCACCTTCC	TGAGAGAAGA	GCACATTCCC	AATGGGACCC	TGGGCTCCAG
CCCTGCCCCA	GCTTGTTGGA	CTAACTCTGG	TGCCCTGCAG	GTCGGCTCCC	AGTGGGCCTC
CCCGGAGAAC	CCCTGCCTCA	TCAATGAGTG	TGTCCGAGTG	AAGGAACAGC	TCTTTATACA
ACAAAAGGAAC	GTCTCCTGCC	CCCAGCTGGA	GGTCCCTGTC	TGCCCCTCGG	GCTTTCAGCT
GAGCTGTAAG	ACCTCAGCGT	GCTGCCCAAG	CTGTCTGCTGT	GGTAAGGCAT	GCAGGCTGGG
GCTGGGCTGG	ACCGGGCACC	ACCTTTAAGC	CTCTCTTTCC	ACTTTTGGCT	CCTGAATTCT
TTGCCTTTTT	AGCAACTTAG	GGAAATGGAT	AGGACCGAAA	TCTTCCCATC	TGCTCCTACT

E46/47

GGAGGCCGCT	GTGGGCTCCA	GCAGTAGGAC	CCCCAGCGCT	GGGTTGTGGG	GTGGGGGGAA
AGGGCCGACC	GATACAGGAG	GGAGGCCCAG	ACACGGAGGA	GGAGCCCCAA	AGAGAGCAGC
CTGCTCGCCG	GTCTCACCAG	GGTGTGTTTT	GCCCACTCTC	ACTCTGCACT	TTTCTCTCCC
CCAGAGCGCA	TGGAGGCCTG	CATGCTCAAT	GGCACTGTCA	TTGGGGTGAG	CCGCTGTCTT
CTTCTCCAGA	GCAAGTGGTG	GGGACAGGGA	AGGGGGTACT	GTGGGAAGGG	GAGCAGGCAA
GTCATTGTAA	AGCAGAAATG	AAGGAAACCA	GAGAGACCCA	ACCCCAGCTT	TCCACTGCCT
GTGGGACGTG	CCTGGCATCA	TGGAGCCCAG	GCTAGGACCA	TCTTCCTGAC	TCTCCGGGCC
TGTCTCACAC	TCACTTCCTG	GCCCCACCT	CAGGCACCTG	TGCATTTCTT	CTGTGTGCAG
AGAAGCACTC	TGAAGTCATT	GTGCACGTTT	TAGTTTGTCC	CCTCTGCCAC	TACCTGGGCT
GCCTCTTTGG	CATGAAAGTT	CTCACTCTTA	CCATCTCGAT	ACTGGAGGTG	GGAGGACGGG
AAGGCAGTGG	GCCATAGGAG	ACAGGAGGAG	CAGCAGAGCG	ATGGCTCATG	GGAGCTATGG
GTGGGTGGGC	AGGAGACAGG	GTATGAGAGT	GAGGTGAGTG	GGGGGTGGG	GGATGCTGGG
GGGCCTGACC	CTGGTGCCTC	TGCTTCCAGC	CCGGGAAGAC	TGTGATGATC	SATGTGAGCA
CGACCTGTCG	CTGCATGGTG	CAGGTGGGGG	TCATCTCTGG	ATTCAAGCTG	SAGTGCAGCA
AGACCACCTG	CAACCCCTGC	CCCCTGGTAA	GAGAGGCTCA	ATGGGGACCG	AGGGCATGGA
CTGGACGCGT	GTGGGACCCA	GGCAGTGGGA	CCTCACTGCG	GTCTTTAAAT	AAATAAATCT

E48

CTACCTAGGG	GACTGAGAAC	TGACAAAAGC	TGGTTGGAGT	TGGATGTTGA	CTTTCCGAAT
TCTAGGTGAA	ACTACTGGAT	GAGAGAGATG	AGAGGCCAGC	AAAATCAGCC	TACTTACAAT
TGGGATGTTT	TAAGGAGGTT	AACATATGGCT	GCTTTTTCCC	CCTCTGGATG	CAGGGTTACA
AGGAAGAAAA	TAACACAGGT	GAATGTTGTG	GCACATGTTT	GCCTACGGCT	TGCACCATTC
AGCTAAGAGG	AGGACAGATC	ATGACACTGA	AGGTAGGAGC	AAGCTGAATG	CAGGGCTCCC
TCACATATCA	CCATCTTGCT	TCCTTTTTTG	GAAATTTTTT	TTCAAGGAAG	AGGGGCTAAG

E49/50

TGGGTGTGGG	GGCTTTATTA	TACACTGTCT	TCTTGTTTCA	TGGTTCTGCA	GATTGTTTCG
TCATCTCCAT	GGAGTCAAGC	TCATGGTTTG	AAGTGGCTTT	GTGAACCAAA	CACTGTCTCT
GACTTTACCC	ACTCCTCTTT	CCTTCCAGCG	TGATGAGACG	CTCCAGGATG	GCTGTGATAC
TCACTTCTG	AAGCTCAATG	AGAGAGGAGA	GTACTTCTGG	GAGAAGAGGG	TCACAGGCTG
CCCACCCTTT	GATGAACACA	AGTGTCTGGC	TGAGGGAGTG	AGTACTCATT	CTGCTTTCCT
CTTTACTGTC	TCAATCTCTA	AGAACAAGAT	TCTTCTGAAT	AGTGTGCATC	CCAGCTCCGG
CATAATTTCT	CAGTGTCTGA	GGCACATCTC	TGGCCTAGTT	GAAGAATCAG	TGAGATTACG
AAATCAAAGC	CTACGGAGAG	ATAAAATTCT	CTGCAATATA	GGATGTTTTA	AAAAAATATT
TTCCTTAAGG	GAGGCTGAGT	GTGTGATTCT	TGAATAAAAT	GTGAGATCAA	ACTGATTTTT
AGTCTCCCTG	GGAATGAAGA	TCCTGATGAC	TCAACTGGAA	GAGAATTCAG	AATCATCAAA
ATTGTTTCAG	CCTGGTCAGG	TAGGGTGGTC	AAGCTGCTCA	CATTTATGAT	AGGAAAGCAT
AGTTTACATC	TGGGCACTTA	AGCACAGGGC	TGTGAGTTCT	GAGCTAAAAA	TTGGCCCCGA
GTGACCTGAA	AGCTGTCTAC	TCTGTACTGT	TTGTGCCTAA	CCTGAAATTT	TGCTGTTTTT
TTAGGGTAAA	ATTATGAAAA	TTCCAGGCAC	CTGCCTCTAC	CCATGTGAGT	GCGTTACTAA
TATCTTGTCC	CTTGAAACCC	ATCAGAGCAA	GTCCAGGGGC	TCTTTGCAGC	TTGCTCCTTG

E51

CTGCATATTT	CGGGCCAGGA	ACAGCTGAGG	TGAACCAGGA	TAAGTGAATT	CCCTCCCTGT
GGCTGGCTTT	ATTTGGTTAC	TGTGAAAGGG	ACCTATTTCC	AGCCAGTGA	GGGCACTGGG
GCTGAAGAGT	GTTCTCTAGA	ACCCAGCCA	GTCCTCAAGT	TTCATCTCTC	ACCTGTCCTG
TAGGTGAGGA	GCCTGAGTGC	AACGACATCA	CTGCCAGGCT	GCAGTATGTC	AAGGTGGGAA
GCTGTAAGTC	TGAAGTAGAG	GTGGATATCC	ACTACTGCCA	GGTAAGGGCT	CTGCTTCAAT
AAGGGCTGGG	TGTGGAGGGC	TGAGCCTCCG	TGTTCCGGATA	CCTATCCTTA	GTTACATTCT
AGAAGGATCT	GGAAAATTCT	CAGGGAAGAG	AAGGGAATA	AACTGGAAGC	ATTTTTTTTT

E52

GATCTTCCTT	TACAAATTGC	CTAGCTGTCC	TGGGTGCTTC	TGGGTGAGAT	CAGACCTGCC
TTGCTTGAG	GGGGTCAGGG	AGAAAGCAGG	CTGCCCCAGA	GCCCTGCCCTA	AGCCAGGACT
TCCCACCATT	GTGAAGCTCC	CATCTTCCTC	TGCTTTCTTG	CAGGGCAAAT	GTGCCAGCAA
AGCCATGTAC	TCCATTGACA	TCAACGATGT	GCAGGACCAG	TGCTCCTGCT	GCTCTCCGAC
ACGGACGGAG	CCCATGCAGG	TGGCCCTGCA	CTGCACCAAT	GGCTCTGTTG	TGTACCATGA
GGTTCTCAAT	GCCATGGAGT	GCAAAATGCTC	CCCCAGGAAG	TGCAGCAAGT	GAGGCTGCTG
CAGCTGCATG	GGTGCCTGCT	GCTGCCTGCC	TTGGCCTGAT	GGCCAGGCCA	GAGTGCTGCC
AGTCCTCTGC	ATGTTCTGCT	CTTGTGCCCT	TCTGAGCCCA	CAATAAAGGC	TGAGCTCTTA
TCTTGC AAAA	GGCTGCTGGT	GCACTGTGTC	GTAGGGCTGA	GATGGCAACG	GTGGGCAGGG
GCTGAGTTCT	GAGACCGGTT	CTGAGGAAGG	AAGCACAGGC	CCCATCTGCA	AGCCTCAGTG
TCAGTGTGAG	ATACTGCCAA	CTTCTGGGGA	AGAGTGATGT	CTTTGCCTCC	

Appendix 4: F8 sequence with probe locations

Partial sequence (20 nt adjacent to ligation site) of the *F8* MLPA probes are indicated for the 26 exons of *F8*.

NM_000132.3 (ENST00000360256) NP_000123.1 (ENSP00000353393) This transcript is a product of gene [ENSG00000185010](#) This transcript is a member of the Human CCDS set: [CCDS35457](#) This transcript corresponds to the following database identifiers: Transcript having exact match between ENSEMBL and HAVANA : [OTTHUMT00000058869](#)

Key

Exons-alternating text colour –blue/black

Codons – alternating background colour

Synonymous SNP –highlighted in green

Non-synonymous SNP – highlighted in yellow

MLPA probes adjacent to ligation site – highlighted in purple and blue

Ambiguity code – underlined letter

```
1  GCTTAGTGCTGAGCACATCCAGTGGTAAAGTTCCTTAAAATGCTCTGCAAAGAAATTGG
   .....
   .....

61  GACTTTTCATTAAATCAGAAATTTACTTTTTCCCCTCCTGGGAGCTAAAGATATTTTA
   .....
   .....

121  GAGAAGAATTAACCTTTTGCTTCTCCAGTTGAACATTTGTAGCAATAAGTCATGCAAATA
   .....ATGCAAATA
   .....M--Q--I-
                                     Exon 1

181  GAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGCATTCTGCTTTTAGTGCCACCAGAAGA
10  GAGCTCTCCACCTGCTTCTTTCTGTGCCTTTTGCATTCTGCTTTTAGTGCCACCAGAAGA
4   -E--L--S--T--C--F--F--L--C--L--L--R--F--C--F--S--A--T--R--R-

241  TACTACCTGGGTGCAGTGGAACTGTCATGGGATATATGCAAAGTGATCTCGGTGAGCTG
70  TACTACCTGGGTGCAGTGGAACTGTCATGGGATATATGCAAAGTGATCTCGGTGAGCTG
24  -Y--Y--L--G--A--V--E--L--S--W--D--Y--M--Q--S--D--L--G--E--L-
                                     Y

301  CCTGTGACGCAAGATTTTCCTCCTAGAGTGCCAAATCTTTTCCATTCAACACCTCAGTC
130  CCTGTGACGCAAGATTTTCCTCCTAGAGTGCCAAATCTTTTCCATTCAACACCTCAGTC
44  -P--V--D--A--R--F--P--P--R--V--P--K--S--F--P--F--N--T--S--V-
                                     Exon 2

361  GTGTACAAAAGACTCTGTTTGTAGAATTCACGGATCACCCTTTCAACATCGCTAAGCCA
190  GTGTACAAAAGACTCTGTTTGTAGAATTCACGGATCACCCTTTCAACATCGCTAAGCCA
64  -V--Y--K--K--T--L--F--V--E--F--T--D--H--L--F--N--I--A--K--P-
                                     W
```

Exon 3

421 AGGCCACCCTGGATGGTCTGCTAGGTCCTACCATCCAGGCTGAGGTTTATGATACAGTG
250 AGGCCACCCTGGATGGTCTGCTAGGTCCTACCATCCAGGCTGAGGTTTATGATACAGTG
84 -R--P--P--W--M--G--L--L--G--P--T--I--Q--A--E--V--Y--D--T--V--

Exon 4

481 GTCATTACACTTAAGAACAATGGCTTCCCATCCTGTCAGTCTTCATGCTGTTGGTGTATCC
310 GTCATTACACTTAAGAACAATGGCTTCCCATCCTGTCAGTCTTCATGCTGTTGGTGTATCC
104 =V--I--T--L--K--N--M--A--S--H--P--V--S--L--H--A--V--G--V--S--

Exon 5

541 TACTGGAAAGCTTCTGAGGGAGCTGAATATGATGATCAGACCAGTCAAAGGGAGAAAGAA
370 TACTGGAAAGCTTCTGAGGGAGCTGAATATGATGATCAGACCAGTCAAAGGGAGAAAGAA
124 -Y--W--K--A--S--E--G--A--E--Y--D--D--Q--T--S--Q--R--E--K--E--

Exon 6

601 GATGATAAAGTCTTCCCTGGTGGAGCCATACATATGTCTGGCAGGTCCTGAAAGAGAAT
430 GATGATAAAGTCTTCCCTGGTGGAGCCATACATATGTCTGGCAGGTCCTGAAAGAGAAT
144 -D--D--K--V--F--P--G--G--S--H--T--Y--V--W--Q--V--L--K--E--N--

Exon 7

661 GGTCCAATGGCCTCTGACCCACTGTGCCTTACCTACTCATATCTTTCTCATGTGGACCTG
490 GGTCCAATGGCCTCTGACCCACTGTGCCTTACCTACTCATATCTTTCTCATGTGGACCTG
164 =G==P--M--A--S--D--P--L--C--L--T--Y--S--Y--L--S--H--V--D--L--

Exon 8

721 GTAAAAGACTTGAATTCAGGCCTCATTGGAGCCCTACTAGTATGTAGAGAAGGAGTCTG
550 GTAAAAGACTTGAATTCAGGCCTCATTGGAGCCCTACTAGTATGTAGAGAAGGAGTCTG
184 -V--K--D--L--N--S--G--L--I--G--A--L--L--V--C--R--E--G--S--L--

Exon 9

781 GCCAAGGAAAAAGACACAGACCTTGCACAAATTTATACTACTTTTGTGCTGTATTTGATGAA
610 GCCAAGGAAAAAGACACAGACCTTGCACAAATTTATACTACTTTTGTGCTGTATTTGATGAA
204 -A--K--E--K--T--Q--T--L--H--K--F--I--L--L--F--A--V--F--D--E--

Exon 10

841 GGGAAAAGTTGGCACTCAGAAACAAAGAACTCCTTGATGCAGGATAGGGATGCTGCATCT
670 GGGAAAAGTTGGCACTCAGAAACAAAGAACTCCTTGATGCAGGATAGGGATGCTGCATCT
224 -G--K--S--W--H--S--E--T--K--N--S--L--M--Q--D--R--D--A--A--S--

Exon 11

901 GCTCGGGCCTGGCCTAAAATGCACACAGTCAATGGTTATGTAAACAGGTCTCTGCCAGGT
730 GCTCGGGCCTGGCCTAAAATGCACACAGTCAATGGTTATGTAAACAGGTCTCTGCCAGGT
244 -A--R--A--W--P--K--M--H--T--V--N--G--Y--V--N--R--S--L--P--G--

Exon 12

961 CTGATTGGATGCCACAGGAAATCAGTCTATTGGCATGTGATTGGAATGGGCACCACTCCT
790 CTGATTGGATGCCACAGGAAATCAGTCTATTGGCATGTGATTGGAATGGGCACCACTCCT
264 -L--I--G--C--H--R--K--S--V--Y--W--H--V--I--G--M--G--T--T--P--

Exon 13

1021 GAAGTGCACTCAATATTCCTCGAAGGTCACACATTTCTTGTGAGGAACCATCGCCAGGCG
850 GAAGTGCACTCAATATTCCTCGAAGGTCACACATTTCTTGTGAGGAACCATCGCCAGGCG
284 -E--V--H--S--I--F--L--E--G--H--T--F--L--V--R--N--H--R--Q--A--

1081 TCCTTGAAATCT^YCGCCAATAACTTTCCTT^RACTGCTCAAACTCTTGATGGACCTTGGA
 910 TCCTTGAAATCTCGCCAATAACTTTCCTTACTGCTCAAACACTCTTGATGGACCTTGGA
 304 -S--L--E--I--=S--P--I--T--F--L--=T--A--Q--T--L--L--M--D--L--G--

Exon 8

1141 CAGTTTCTAC^YTGTTT^VGTTCATATCTCTTCCACCAACATGATGGCATGGAAGCTTATGTC
 970 CAGTTTCTACTGTTTGTTCATATCTCTTCCACCAACATGATGGCATGGAAGCTTATGTC
 324 -Q--F--L--=L--F--=C--H--I--S--S--H--Q--H--D--G--M--E--A--Y--V--

1201 AAAGTAGACAGCTGTCCAGAGGAACCCCAACTACGAATGAAAAATAATGAAGAAGC^RGGAA
 1030 AAAGTAGACAGCTGTCCAGAGGAACCCCAACTACGAATGAAAAATAATGAAGAAGCGGAA
 344 -K--V--D--S--C--P--E--E--P--Q--L--R--M--K--N--N--E--E--A--E--

1261 GACTATGATGATGATCTTACTGATTCTGAAATGGATGTGGTCAGGTTTGATGATGACAAC
 1090 GACTATGATGATGATCTTACTGATTCTGAAATGGATGTGGTCAGGTTTGATGATGACAAC
 364 -D--Y--D--D--D--L--T--D--S--E--M--D--V--V--R--F--D--D--D--N--

1321 TCTCCTTCCTTTATCCAAATTC^RCTCAGTTGCCAAGAAGCATCCTAAAAC^{YY}TGGGTACAT
 1150 TCTCCTTCCTTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCCTAAAAC^RTGGGTACAT
 384 -S--P--S--F--I--Q--I--=R--S--V--A--K--K--H--P--K--T--W--V--H--

1381 TACAT^KTGCTGCTGAAG^RAGGAGGACTGGGACTATGCTCCCTTAGTCCTCGCCCCGATGAC
 1210 TACATTGCTGCTGAAGAGGAGGACTGGGACTATGCTCCCTTAGTCCTCGCCCCGATGAC
 404 -Y--=I--A--A--E--=E--E--D--W--D--Y--A--P--L--V--L--A--P--D--D--

Exon 9

1441 AGAAGTTATAAAAGTCAATATTT^KGAACAATGGCCCTCAGCGGATTGGTAGGAAGTACAAA
 1270 AGAAGTTATAAAAGTCAATATTTGAACAATGGCCCTCAGCGGATTGGTAGGAAGTACAAA
 424 -R--S--Y--K--S--Q--Y--=L--N--N--G--P--Q--R--I--G--R--K--Y--K--

1501 AAAGTCCGATTTATGGCA^RTACACAGATGAAACCTTTAAGACTCGTGAAGCTATT^WCAGCAT
 1330 AAAGTCCGATTTATGGCATAACACAGATGAAACCTTTAAGACTCGTGAAGCTATT^RCAGCAT
 444 =K--V--R--F--M--A--=Y--T--D--E--T--F--K--T--R--E--A--I--Q--H--

Exon 10

1561 GAATCAGGAATCTTGGGACCTTTACTTTATGGGGAAGTTGGAGACACACTGTTGATTATA
 1390 GAATCAGGAATCTTGGGACCTTTACTTTATGGGGAAGTTGGAGACACACTGTTGATTATA
 464 -E--S--G--I--L--G--P--L--L--Y--G--E--V--G--D--T--L--L--I--I--

1621 TTTAAGAATCAAGCAAGCAGACCATAT^{YR}TACCTTACCCT^YCACGGAATCACTGATGTCC^RGT
 1450 TTTAAGAATCAAGCAAGCAGACCATATAACATCTACCCTCACGGAATCACTGATGTCCGT
 484 -F--K--N--Q--A--S--R--P--=Y--N--=I--Y--P--H--=G--I--T--D--V--=R--

Exon 11

1681 CCTTTGTATTCAAGGAGATTACCAAAAG^RGTGTAACCATTTGAAGGATTTTCCAATTCTG
 1510 CCTTTGTATTCAAGGAGATTACCAAAAG^RGTGTAACCATTTGAAGGATTTTCCAATTCTG
 504 -P--L--Y--S--R--R--L--P--K--G--V--K--H--L--K--D--F--P--I--L--

1741 CCAGGAGAAATATTCAAATATAAAATGGACAGTGACTGTAGAAGATGGGCCAACTAAATCA
 1570 CCAGGAGAAATATTCAAATATAAAATGGACAGTGACTGTAGAAGATGGGCCAACTAAATCA
 524 -P--G--E--I--F--K--Y--K--W--T--V--T--V--E--D--G--P--T--K--S-

R Y BR R R
 1801 GATCCTCGGTGCCTGACC CGCTATTACTCTAGTTTCGTTAATATGGAGAGAGATCTAGCT
 1630 GATCCTCGGTGCCTGACCCGCTATTACTCTAGTTTCGTTAATATGGAGAGAGATCTAGCT
 544 =D=-P=-R=-C--L--T=-R=-Y--Y--S=-S=-F--V--N--M--E--R=-D=-L--A-

YY
 1861 TCAGGACTCATTGGCCCTCTCCTCATCTGCTACAAAGAATCTGTAGATCAAAGAGGAAAC
 1690 TCAGGACTCATTGGCCCTCTCCTCATCTGCTACAAAGAATCTGTAGATCAAAGAGGAAAC
 564 -S--G--L--I--G--P--L--L--I--C--Y--K--E--S=-V--D--Q--R--G--N-

Exon 12

1921 CAGATAATGTCAGACAAGAGGAATGTCATCCTGTTTTCTGTATTTGATGAGAAC GAAGC
 1750 CAGATAATGTCAGACAAGAGGAATGTCATCCTGTTTTCTGTATTTGATGAGAACCGAAGC
 584 -Q--I--M--S--D--K--R--N--V--I--L--F--S--V--F--D--E--N--R--S-

Y
 1981 TGGTACCTCACAGAGAATATACAACGCTTTCTCCCAATCCAGCTGGAGTGCAGCTTGAG
 1810 TGGTACCTCACAGAGAATATACAACGCTTTCTCCCAATCCAGCTGGAGTGCAGCTTGAG
 604 -W--Y--L--T--E--N--I--Q--R--F--L--P--N--P--A--G--V--Q--L--E-

Exon 13

2041 GATCCAGAGTTCCAAGCCTCCAACATCATGCACAGCATCAATGGCTATGTTTTTGATAGT
 1870 GATCCAGAGTTCCAAGCCTCCAACATCATGCACAGCATCAATGGCTATGTTTTTGATAGT
 624 -D--P--E--F--Q--A--S--N--I--M--H--S--I--N--G--Y--V--F--D--S-

2101 TTGCAGTTGTCAGTTTGTTTGCATGAGGTGGCATACTGGTACATTCTAAGCATTGGAGCA
 1930 TTGCAGTTGTCAGTTTGTTTGCATGAGGTGGCATACTGGTACATTCTAAGCATTGGAGCA
 644 -L--Q--L--S--V--C--L--H--E--V--A--Y--W--Y--I--L--S--I--G--A-

2161 CAGACTGACTTCCTTTCTGTCTTCTTCTCTGGATATACCTTCAAACACAAAATGGTCTAT
 1990 CAGACTGACTTCCTTTCTGTCTTCTTCTCTGGATATACCTTCAAACACAAAATGGTCTAT
 664 -Q--T--D--F--L--S--V--F--F--S--G--Y--T--F--K--H--K--M--V--Y-

2221 GAAGACACACTCACCCCTATTCCCATTTCTCAGGAGAACTGTCTTCATGTTCGATGGAAAAC
 2050 GAAGACACACTCACCCCTATTCCCATTTCTCAGGAGAACTGTCTTCATGTTCGATGGAAAAC
 684 -E--D--T--L--T--L--F--P--F--S--G--E--T--V--F--M--S--M--E--N-

Exon 14A

2281 CCAGGTCTATGGATTCTGGGGTGCCACAACTCAGACTTTTCGGAACAGAGGCATGACCGCC
 2110 CCAGGTCTATGGATTCTGGGGTGCCACAACTCAGACTTTTCGGAACAGAGGCATGACCGCC
 704 -P--G--L--W--I--L--G--C--H--N--S--D--F--R--N--R--G--M--T--A-

R
 2341 TTAAGTGAAGGTTTCTAGTTGTGACAAGAACAAGTGGTGAATATTACGAGGACAGTTATGAA
 2170 TTAAGTGAAGGTTTCTAGTTGTGACAAGAACAAGTGGTGAATATTACGAGGACAGTTATGAA
 724 -L--L--K--V--S--S--C--D--K--N--T--G--D--Y--Y--E=-D--S--Y--E-

2401 GATATTTTCAGCATACTTGCTGAGTAAAAACAATGCCATTGAACCAAGAAGCTTCTCCCAG
 2230 GATATTTTCAGCATACTTGCTGAGTAAAAACAATGCCATTGAACCAAGAAGCTTCTCCCAG
 744 -D--I--S--A--Y--L--L--S--K--N--N--A--I--E--P--R--S--F--S--Q-

2461 AATTCAAGACACCTTAGCACTAGGCAAAAGCAATTTAATGCCACCACAATTCCAGAAAAT
 2290 AATTCAAGACACCTTAGCACTAGGCAAAAGCAATTTAATGCCACCACAATTCCAGAAAAT
 764 -N--S--R--H--P--S--T--R--Q--K--Q--F--N--A--T--T--I--P--E--N-

2521 GACATAGAGAAGACTGACCCTTGGTTTGCACACA^RGAACACCTATGCCTAAAATACAAAAT
 2350 GACATAGAGAAGACTGACCCTTGGTTTGCACACAGAACACCTATGCCTAAAATACAAAAT
 784 -D--I--E--K--T--D--P--W--F--A--H--R--T--P--M--P--K--I--Q--N-

2581 GTCTCCTCTAGTGATTTGTTGATGCTCTTGCAGACAGAGTCCTACTCCACATGGGCTATCC
 2410 GTCTCCTCTAGTGATTTGTTGATGCTCTTGCAGACAGAGTCCTACTCCACATGGGCTATCC
 804 -V--S--S--S--D--L--L--M--L--L--R--Q--S--P--T--P--H--G--L--S-

2641 TTA^{14BF}TCTGATCTCCA^{GAAGCCAAATATGAGACTTTTTCTGATGATCCATCACCTGGAGCA}
 2470 TTATCTGATCTCCAAGAAGCCAAATATGAGACTTTTTCTGATGATCCATCACCTGGAGCA
 824 -L--S--D--L--Q--E--A--K--Y--E--T--F--S--D--D--P--S--P--G--A-

2701 ATAGACAGTAATAACAGCCTGTCTGAAATGACACACTTCAGGCCACAGCTCCATCACAGT
 2530 ATAGACAGTAATAACAGCCTGTCTGAAATGACACACTTCAGGCCACAGCTCCATCACAGT
 844 -I--D--S--N--N--S--L--S--E--M--T--H--F--R--P--Q--L--H--H--S-

2761 GGGGACATGGTATTTACCCCTGAGTCAGGCCTCCAATTAAGATTAAATGAGAACTGGGG
 2590 GGGGACATGGTATTTACCCCTGAGTCAGGCCTCCAATTAAGATTAAATGAGAACTGGGG
 864 -G--D--M--V--F--T--P--E--S--G--L--Q--L--R--L--N--E--K--L--G-

2821 ACAACTGCAGCAACAGAGTTGAAGAACTTGATTTCAAAGTTTCTAGTACATCAAATAAT
 2650 ACAACTGCAGCAACAGAGTTGAAGAACTTGATTTCAAAGTTTCTAGTACATCAAATAAT
 884 -T--T--A--A--T--E--L--K--K--L--D--F--K--V--S--S--T--S--N--N-

2881 CTGATTTCAACAATTCCATCAGACAATTTGGCAGCAGGTACTGATAATACAAGTTCCTTA
 2710 CTGATTTCAACAATTCCATCAGACAATTTGGCAGCAGGTACTGATAATACAAGTTCCTTA
 904 -L--I--S--T--I--P--S--D--N--L--A--A--G--T--D--N--T--S--S--L-

2941 GGACCCCAAGTATGCCAGTTCATTATGATAGTCAATTAGATACCACTCTATTTGGCAAA
 2770 GGACCCCAAGTATGCCAGTTCATTATGATAGTCAATTAGATACCACTCTATTTGGCAAA
 924 -G--P--P--S--M--P--V--H--Y--D--S--Q--L--D--T--T--L--F--G--K-

3001 AAGTCATCTCCCCTTACTGAGTCTGGTGGACCTCTGAGCTTGAGTGAAGAAAATAATGAT
 2830 AAGTCATCTCCCCTTACTGAGTCTGGTGGACCTCTGAGCTTGAGTGAAGAAAATAATGAT
 944 -K--S--S--P--L--T--E--S--G--G--P--L--S--L--S--E--E--N--N--D-

3061 TCAAAGTTGTTAGAAATCAGGTTTAATGAATAGCCAAGAAAGTTCATGGGGAAAAAAT^RGTA
 2890 TCAAAGTTGTTAGAAATCAGGTTTAATGAATAGCCAAGAAAGTTCATGGGGAAAAAATGTA
 964 -S--K--L--L--E--S--G--L--M--N--S--Q--E--S--S--W--G--K--N--=V=

3121 TCGTCAACAGAGAGTGGTAGGTTATTTAAAGGGAAAAGAGCTCATGGACCTGCTTTGTTG
 2950 TCGTCAACAGAGAGTGGTAGGTTATTTAAAGGGAAAAGAGCTCATGGACCTGCTTTGTTG
 984 -S--S--T--E--S--G--R--L--F--K--G--K--R--A--H--G--P--A--L--L--

3181 ACTAAAGATAATGCCTTATTCAAAGTTAGCATCTCTTTGTTAAAGACAAACAAAACCTTCC
 3010 ACTAAAGATAATGCCTTATTCAAAGTTAGCATCTCTTTGTTAAAGACAAACAAAACCTTCC
 1004 -T--K--D--N--A--L--F--K--V--S--I--S--L--L--K--T--N--K--T--S--

3241 AATAATTCAGCAACTAATAGAAAGACTCACATTGATGGCCCATCATTATTAATTGAGAAT
 3070 AATAATTCAGCAACTAATAGAAAGACTCACATTGATGGCCCATCATTATTAATTGAGAAT
 1024 -N--N--S--A--T--N--R--K--T--H--I--D--G--P--S--L--L--I--E--N--

3301 AGTCCATCAGTCTGGCAAATATATTAGAAAGTGACACT^RGAGTTTAAAAAAGTGACACCT
 3130 AGTCCATCAGTCTGGCAAATATATTAGAAAGTGACACTGAGTTTAAAAAAGTGACACCT
 1044 -S--P--S--V--W--Q--N--I--L--E--S--D--T--=E--F--K--K--V--T--P--

3361 TTGATTCATGACAGAATGCTTATGGACAAAAATGCTACAGCTTTGAGGCTAAATCATATG
 3190 TTGATTCATGACAGAATGCTTATGGACAAAAATGCTACAGCTTTGAGGCTAAATCATATG
 1064 -L--I--H--D--R--M--L--M--D--K--N--A--T--A--L--R--L--N--H--M--

3421 TCAAATAAAACTACTTCATCAAAAAACATGGAAATGGTCCAACAGAAAAAAGAGGGCCCC
 3250 TCAAATAAAACTACTTCATCAAAAAACATGGAAATGGTCCAACAGAAAAAAGAGGGCCCC
 1084 -S--N--K--T--T--S--S--K--N--M--E--M--V--Q--Q--K--K--E--G--P--

3481 ATTCACCAGATGCACAAAATCCAGATATGTCGTTCTTTAAGATGCTATTCTTGCCAGAA
 3310 ATTCACCAGATGCACAAAATCCAGATATGTCGTTCTTTAAGATGCTATTCTTGCCAGAA
 1104 -I--P--P--D--A--Q--N--P--D--M--S--F--F--K--M--L--F--L--P--E--

3541 TCAGCAAGGTGGATACAAAGGACTCATGGAAAGAACTCTCTGAACTCTGGCAAGGCCCC
 3370 TCAGCAAGGTGGATACAAAGGACTCATGGAAAGAACTCTCTGAACTCTGGGCAAGGCCCC
 1124 -S--A--R--W--I--Q--R--T--H--G--K--N--S--L--N--S--G--Q--G--P--

3601 AGTCCAAAGCAATTAGTATCCTTAGGACCAGAAAAATCTGTGGAAGGTCAGAATTTCTTG
 3430 AGTCCAAAGCAATTAGTATCCTTAGGACCAGAAAAATCTGTGGAAGGTCAGAATTTCTTG
 1144 -S--P--K--Q--L--V--S--L--G--P--E--K--S--V--E--G--Q--N--F--L--

3661 TCTGAGAAAAACAAAGTGGTAGTAGGAAAGGGTGAATTTACAAAGGACGTAGGACTCAAA
 3490 TCTGAGAAAAACAAAGTGGTAGTAGGAAAGGGTGAATTTACAAAGGACGTAGGACTCAAA
 1164 -S--E--K--N--K--V--V--V--G--K--G--E--F--T--K--D--V--G--L--K--

14DF
 3721 GAGATGGTTTTTCCAAGCAGCAGAAACCTATTTCTTACTAACTTGGATAATTTACATGAA
 3550 GAGATGGTTTTTCCAAGCAGCAGAAACCTATTTCTTACTAACTTGGATAATTTACATGAA
 1184 -E--M--V--F--P--S--S--R--N--L--F--L--T--N--L--D--N--L--H--E-

Y
 3781 AATAATACACA^YCAATCAAGAAAAAAAAAATTCAGGAAGAAATAGAAAAGAAGGAAACATTA
 3610 AATAATACACACAATCAAGAAAAAAAAAATTCAGGAAGAAATAGAAAAGAAGGAAACATTA
 1204 -N--N--T--H--N--Q--E--K--K--I--Q--E--E--I--E--K--K--E--T--L-

14CR
 3841 ATCCAAGAGAATGTAGTTTTGCCTCAGATACATACAGTGACTGGCACTAAGAATTTTCATG
 3670 ATCCAAGAGAATGTAGTTTTGCCTCAGATACATACAGTGACTGGCACTAAGAATTTTCATG
 1224 -I--Q--E--N--V--V--L--P--Q--I--H--T--V--T--G--T--K--N--F--M-

S
 3901 AAGAACCCTTTTCTTACTGAGCACTAGGCCAAATGTAGAAGGTTTCATATGA^SGGGGCATAT
 3730 AAGAACCCTTTTCTTACTGAGCACTAGGCCAAATGTAGAAGGTTTCATATGACGGGGCATAT
 1244 -K--N--L--F--L--L--S--T--R--Q--N--V--E--G--S--Y--D--G--A--Y-

3961 GCTCCAGTACTTCAAGATTTTAGGTCATTAAATGATTCAACAAATAGAACAAAGAAACAC
 3790 GCTCCAGTACTTCAAGATTTTAGGTCATTAAATGATTCAACAAATAGAACAAAGAAACAC
 1264 -A--P--V--L--Q--D--F--R--S--L--N--D--S--T--N--R--T--K--K--H-

M
 4021 ACAGCTCATTTCTC^MAAAAAAGGGGAGGAAGAAAACCTTGAAGGCTTGGGAAATCAAACC
 3850 ACAGCTCATTTCTCAAAAAAGGGGAGGAAGAAAACCTTGAAGGCTTGGGAAATCAAACC
 1284 -T--A--H--F--S--K--K--G--E--E--E--N--L--E--G--L--G--N--Q--T-

M
 4081 AAGCAAATTGT^MAGAGAAATATGCATGCACCACAAGGATATCTCCTAATACAAGCCAGCAG
 3910 AAGCAAATTGTAGAGAAATATGCATGCACCACAAGGATATCTCCTAATACAAGCCAGCAG
 1304 -K--Q--I--V--E--K--Y--A--C--T--T--R--I--S--P--N--T--S--Q--Q-

R
 4141 AATTTTGTACGCAACGTAGTAAG^RAGAGCTTTGAAACAATTCAGACTCCCACTAGAAGAA
 3970 AATTTTGTACGCAACGTAGTAAGAGAGCTTTGAAACAATTCAGACTCCCACTAGAAGAA
 1324 -N--F--V--T--Q--R--S--K--R--A--L--K--Q--F--R--L--P--L--E--E-

14EF
 4201 ACAGAACTTGAAAAAAGGATAATTGTGGATGACACCTCAACCCAGTGGTCCAAAAACATG
 4030 ACAGAACTTGAAAAAAGGATAATTGTGGATGACACCTCAACCCAGTGGTCCAAAAACATG
 1344 -T--E--L--E--K--R--I--I--V--D--D--T--S--T--Q--W--S--K--N--M-

4261 AAACATTTGACCCCGAGCACCTCACACAGATAGACTACAATGAGAAGGAGAAAGGGGCC
 4090 AAACATTTGACCCCGAGCACCTCACACAGATAGACTACAATGAGAAGGAGAAAGGGGCC
 1364 -K--H--L--T--P--S--T--L--T--Q--I--D--Y--N--E--K--E--K--G--A-

14DR

4321 ATTACTCAGTCTCCCTTATCAGATTGCCTTACGAGGAGTCATAGCATCCCTCAAGCAAAT
4150 ATTACTCAGTCTCCCTTATCAGATTGCCTTACGAGGAGTCATAGCATCCCTCAAGCAAAT
1384 -I--T--Q--S--P--L--S--D--C--L--T--R--S--H--S--I--P--Q--A--N--

4381 AGATCTCCATTACCCATTGCAAAGGTATCATCATTTCCATCTATTAGACCTATATATCTG
4210 AGATCTCCATTACCCATTGCAAAGGTATCATCATTTCCATCTATTAGACCTATATATCTG
1404 -R--S--P--L--P--I--A--K--V--S--S--F--P--S--I--R--P--I--Y--L--

4441 ACCAGGGTCCTATTCCAAGACAACCTCTTCTCATCTTCCAGCAGCATCTTATAGAAAGAAA
4270 ACCAGGGTCCTATTCCAAGACAACCTCTTCTCATCTTCCAGCAGCATCTTATAGAAAGAAA
1424 -T--R--V--L--F--Q--D--N--S--S--H--L--P--A--A--S--Y--R--K--K--

4501 GATTCTGGGGTCCAAGAAAGCAGTCATTTCTTACAAGGAGCCAAAAAAATAACCTTTCT
4330 GATTCTGGGGTCCAAGAAAGCAGTCATTTCTTACAAGGAGCCAAAAAAATAACCTTTCT
1444 -D--S--G--V--Q--E--S--S--H--F--L--Q--G--A--K--K--N--N--L--S--

4561 TTAGCCATTCTAACCTTGGAGATGACTGGTGATCAAAGAGAGGTTGGCTCCCT^YTGGGGACA
4390 TTAGCCATTCTAACCTTGGAGATGACTGGTGATCAAAGAGAGGTTGGCTCCCTGGGGACA
1464 -L--A--I--L--T--L--E--M--T--G--D--Q--R--E--V--G--S--~~L~~--G--T--

4621 AGTGCCACAAATTCAGT^SCACATACAAGAAAGTTGAGAACACTGTTCTCCC^RGAAACCAGAC
4450 AGTGCCACAAATTCAGTCACATACAAGAAAGTTGAGAACACTGTTCTCCCGAAACCAGAC
1484 -S--A--T--N--S--V--T--Y--K--K--V--E--N--T--V--L--P--K--P--D--

4681 TTGCCCCAAAC^MATCTGGCAAAGTTGAATTGCTTCCAAAAGTTCACATTTATCAGAAGGAC
4510 TTGCCCCAAACATCTGGCAAAGTTGAATTGCTTCCAAAAGTTCACATTTATCAGAAGGAC
1504 -L--P--K--T--S--G--K--V--E--L--L--P--K--V--H--I--Y--Q--K--D--

14FF **14ER** ^Y

4741 CTATTCCCTACGGAAACTAGCAATGGGTCTCCTGGCCATCTGGATCTCGTGGAAGGGAGC^G
4570 CTATTCCCTACGGAAACTAGCAATGGGTCTCCTGGCCATCTGGATCTCGTGGAAGGGAGC
1524 -L--E--P--T--E--T--S--N--G--S--P--G--H--L--D--L--V--E--G--S--

4801 CTTCTTCAGGGAACAGAGGGAGCGATTAAGTGAATGAAGCAAACAGACCTGGAAAAGTT
4630 CTTCTTCAGGGAACAGAGGGAGCGATTAAGTGAATGAAGCAAACAGACCTGGAAAAGTT
1544 -L--L--Q--G--T--E--G--A--I--K--W--N--E--A--N--R--P--G--K--V--

4861 CCCTTTCTGAGAGTAGCAACAGAAAGCTCTGCAAAGACTCCCTCCAAGCTATTGGATCCT
4690 CCCTTTCTGAGAGTAGCAACAGAAAGCTCTGCAAAGACTCCCTCCAAGCTATTGGATCCT
1564 -P--F--L--R--V--A--T--E--S--S--A--K--T--P--S--K--L--L--D--P--

4921 CTTGCTTGGGATAACCACTATGGTACTCAGATACCAAAGAAGAGTGGAAATCCCAAGAG
4750 CTTGCTTGGGATAACCACTATGGTACTCAGATACCAAAGAAGAGTGGAAATCCCAAGAG
1584 -L--A--W--D--N--H--Y--G--T--Q--I--P--K--E--E--W--K--S--Q--E--

4981 AAGTCACCAGAAAAACAGCTTTTAAGAAAAAGGATACCATTTTGTCCCTGAACGCTTGT
 4810 AAGTCACCAGAAAAACAGCTTTTAAGAAAAAGGATACCATTTTGTCCCTGAACGCTTGT
 1604 -K--S--P--E--K--T--A--F--K--K--K--D--T--I--L--S--L--N--A--C-

5041 GAAAGCAATCATGCAATAGCAGCAATAAATGAGGGACAAAATAAGCCC^WGAAATAGAAGT
 4870 GAAAGCAATCATGCAATAGCAGCAATAAATGAGGGACAAAATAAGCCC^WGAAATAGAAGTC
 1624 -E--S--N--H--A--I--A--A--I--N--E--G--Q--N--K--P--E--I--E--V-

5101 ACCTGGGCAAAGCAAGGTAGGACTGAAAGGCTGTGCTCTCAA^WAACCCACCAGTCTTGAAA
 4930 ACCTGGGCAAAGCAAGGTAGGACTGAAAGGCTGTGCTCTCAA^WAACCCACCAGTCTTGAAA
 1644 -T--W--A--K--Q--G--R--T--E--R--L--C--S--Q--N--P--P--V--L--K-

5161 CGCCATCAACGGGAAATAACTCGTACTACTCTTCAGTCAGATCAAGAGGAAATTGACTAT
 4990 CGCCATCAACGGGAAATAACTCGTACTACTCTTCAGTCAGATCAAGAGGAAATTGACTAT
 1664 -R--H--Q--R--E--I--T--R--T--T--L--Q--S--D--Q--E--E--I--D--Y-

5221 GATGATACCATATCAGTTGAAATGAAGAAGGAAGATTTTGACATT^WTATGATGAGGATGAA
 5050 GATGATACCATATCAGTTGAAATGAAGAAGGAAGATTTTGACATT^WTATGATGAGGATGAA
 1684 -D--D--T--I--S--V--E--M--K--K--E--D--F--D--I--^WY--D--E--D--E-

5281 AATCAGAGCCCCCGCAGCTTTCAAAGAAAACACGACACTATTTTATTGCTGCAGTGGAG
 5110 AATCAGAGCCCCCGCAGCTTTCAAAGAAAACACGACACTATTTTATTGCTGCAGTGGAG
 1704 -N--Q--S--P--R--S--F--Q--K--K--T--R--H--Y--F--I--A--A--V--E-

Exon 15

5341 AGGCTCTGGGATTATGGGATGAGTAGCTCCCCACATGTTCTAAGAAACAG^WGCTCAGAGT
 5170 AGGCTCTGGGATTATGGGATGAGTAGCTCCCCACATGTTCTAAGAAACAG^WGCTCAGAGT
 1724 -R--L--W--D--Y--G--M--S--S--S--P--H--V--L--R--N--R--^WA--Q--S-

5401 GGCAGTGTCCCTCAGTTCAAGAAAGTTGTTT^WTCCAGGAATTTACTGATGGCTCCTTTACT
 5230 GGCAGTGTCCCTCAGTTCAAGAAAGTTGTTT^WTCCAGGAATTTACTGATGGCTCCTTTACT
 1744 -G--S--V--P--Q--F--K--K--V--V--F--Q--E--F--T--D--G--S--F--T-

5461 CAG^RCCCTTATACCGTGGAGAACTAAATGAACAT^{K S}TTGGGACTCCTGG^RGGCCATATATAAGA
 5290 CAGCCCTTATACCGTGGAGAACTAAATGAACATTTGGGACTCCTGGGGGCCATATATAAGA
 1764 -Q--P--L--Y--^RR--G--E--L--N--E--H--^{K S}L--G--L--L--^RG--P--Y--I--R-

Exon 16

5521 GCAGAA^YGTTGAAGATAATATCATG^YGTAAC^{BR}TTTCAGAAATCAGGCCTCT^MCGTCCCTATTCC
 5350 GCAGAA^YGTTGAAGATAATATCATG^YGTAAC^{BR}TTTCAGAAATCAGGCCTCTCGTCCCTATTCC
 1784 -A--E--V--E--D--N--I--M--^YV--T--F--R--N--Q--A--S--^{BR}R--P--Y--^MS=

5581 TTCTATTCTAGCCTTATTTCTTATGAGGAAGATCAGAGGCAAGGAGCAGAACCTAGAAAA
 5410 TTCTATTCTAGCCTTATTTCTTATGAGGAAGATCAGAGGCAAGGAGCAGAACCTAGAAAA
 1804 -F--Y--S--S--^YL--I--S--Y--E--E--D--Q--R--Q--G--A--E--P--R--K-

5641 AACTTTGTCAAGCCTAATGAAACCAAACTTACTTTTGGAAAGTGCAACATCATATG^RGCA
5470 AACTTTGTCAAGCCTAATGAAACCAAACTTACTTTTGGAAAGTGCAACATCATATGGCA
1824 -N--F--V--K--P--N--E--T--K--T--Y--F--W--K--V--Q--H--H--M--A--

Exon 17

5701 ^YCCC^MACTAAAGATGAGTTTGACTGCAAAG^YCTGGGCTTATTTCTCTGATGTTGACCTGGAA
5530 CCCACTAAAGATGAGTTTGACTGCAAAGCCTGGGCTTATTTCTCTGATGTTGACCTGGAA
1844 =P==T=-K--D--E--F--D--C--K--A=-W--A--Y--F--S--D--V--D--L--E-

5761 ^DAAAGATGTGC^RACTCAGGCCTGATTGGAC^SCCTTCTGGTCTGCCACACTAACACACTGAAC
5590 AAAGATGTGCACTCAGGCCTGATTGGACCCCTTCTGGTCTGCCACACTAACACACTGAAC
1864 -K--D=-V--H=-S--G--L--I--G--P=-L--L--V--C--H--T--N--T--L--N--

5821 CCTGCTCATGGGAGACAAGTGACAGTACAGGAATTTGCTCTGTTTTTCACCATCTTTGAT
5650 CCTGCTCATGGGAGACAAGTGACAGTACAGGAATTTGCTCTGTTTTTCACCATCTTTGAT
1884 -P--A--H--G--R--Q--V--T--V--Q--E--F--A--L--F--F--T--I--F--D--

5881 ^RGAGACCAAAAGCTGGTACTTCACTGAAAATATGAAAGAAACTGCAGGGCTCCCT^{SS}GCAAT
5710 GAGACCAAAAGCTGGTACTTCACTGAAAATATGAAAGAAACTGCAGGGCTCCCTGCAAT
1904 =E=-T--K--S--W--Y--F--T--E--N--M--E--R--N--C--R--A--P--C=-N--

Exon 18

5941 ATCCAGATGGAAGATCCCACTTTTAAAGAGAATTATCGCTTCCATG^RCAATCAATGGCTAC
5770 ATCCAGATGGAAGATCCCACTTTTAAAGAGAATTATCGCTTCCATGCAATCAATGGCTAC
1924 -I--Q--M--E--D--P--T--F--K--E--N--Y--R--F--H--A--I--N=-G--Y--

6001 ATAATGGATACACTACCTGGCTTAGTAATGGCTCAGGATCAAAGGATT^DCGATGGTATCTG
5830 ATAATGGATACACTACCTGGCTTAGTAATGGCTCAGGATCAAAGGATTGATGGTATCTG
1944 -I--M--D--T--L--P--G--L--V--M--A--Q--D--Q--R--I--R=-W--Y--L--

6061 ^RCTCAGCATGGGCAGCAATGAAAACA^KTCCATTCTATTCA^YTTCAGTGACATGTGTTCACT
5890 CTCAGCATGGGCAGCAATGAAAACATCCATTCTATTCA^KTTCAGTGACATGTGTTCACT
1964 -L--S--M--G=-S--N--E--N--I--H--S--I--H--F--S--G==H=-V--F--T--

Exon 19

6121 GTACGAAAAAAGAGGAGTATAAAATGGCACTGTACAATCTCTATCCAGGTGTTTTTGAG
5950 GTACGAAAAAAGAGGAGTATAAAATGGCACTGTACAATCTCTATCCAGGTGTTTTTGAG
1984 -V--R--K--K--E--E--Y--K--M--A--L--Y--N--L--Y--P--G--V--F--E--

6181 ACAGTGGAATGTTACCATCCAAAGCTGGAATTT^YGGCGGGTGAATGCCTTATTGGCGAG
6010 ACAGTGGAATGTTACCATCCAAAGCTGGAATTTGGCGGGTGAATGCCTTATTGGCGAG
2004 -T--V--E--M--L--P--S--K--A--G--I--W--R=-V--E--C--L--I--G--E--

Exon 20

6241 CATCTACATGCTGGGATGAGCACACTTTTTCTGGTGTACAGCAATAAGTGTCTAGACTGGG
6070 CATCTACATGCTGGGATGAGCACACTTTTTCTGGTGTACAGCAATAAGTGTCTAGACTCCC
2024 -H--L--H--A--G--M--S--T--L--F--L--V--Y--S--N--K--C--Q--T--P--

Exon 21

6301 CTGGGAATGGCTTCTGGACACATTAGAGATTTTCAGATTACAGCTTCAGGACAATATGGA
6130 CTGGGAATGGCTTCTGGACACATTAGAGATTTTCAGATTACAGCTTCAGGACAATATGGA
2044 -L--G--M--A--S--G--H--I--R--D--F--Q--I--T--A--S--G--Q--Y--G--

6361 CAG^YTGGGCCCAAAGCTGGCCAGACTTTCATTAT^YTCCGGATCAATCAATGCCTGGAGCACC
6190 CAGTGGGCCCAAAGCTGGCCAGACTTTCATTAT^YTCCGGATCAATCAATGCCTGGAGCACC
2064 -Q--W--A--P--K--L--A--R--L--H--Y--S--G--S--I--N--A--W--S--T--

Exon 22

6421 AAGGAGCCCTTTT^YCTTGGATCAAG^RGTGGATCTGTTGGCACCAATGATTATTCACGGCATC
6250 AAGGAGCCCTTTTCTTGGATCAAG^RGTGGATCTGTTGGCACCAATGATTATTCACGGCATC
2084 -K--E--P--F--S--W--I--K--V--D--L--L--A--P--M--I--I--H--G--I--

6481 AAGACCAGGGTGCCCGTCAGAAGTTCTCCAGCCTCTACATCTCTCAGTTTATCATCATG
6310 AAGACCAGGGTGCCCGTCAGAAGTTCTCCAGCCTCTACATCTCTCAGTTTATCATCATG
2104 -K--T--Q--G--A--R--Q--K--F--S--S--L--Y--I--S--Q--F--I--I--M--

6541 TATAGTCTTGATGGGAAGAAGTGGCAGACTTATCGAGGAAATTCCACTGGAACCTTAATG
6370 TATAGTCTTGATGGGAAGAAGTGGCAGACTTATCGAGGAAATTCCACTGGAACCTTAATG
2124 -Y--S--L--D--G--K--K--W--Q--T--Y--R--G--N--S--T--G--T--L--M--

Exon 23

6601 GTCTTCTTTGGCAATGTGGATTTCATCTGGGATAAAACACAATATTTTAAACCCTCCAATT
6430 GTCTTCTTTGGCAATGTGGATTTCATCTGGGATAAAACACAATATTTTAAACCCTCCAATT
2144 -V--F--F--G--N--V--D--S--S--G--I--K--H--N--I--F--N--P--P--I--

6661 ATTGCTCGATACATCCGTTTGCACCCAACCTCATTTATAGCATT^{YD}CGCAGCACTCTT^{YR}CGCATG
6490 ATTGCTCGATACATCCGTTTGCACCCAACCTCATTTATAGCATT^{YD}CGCAGCACTCTT^{YR}CGCATG
2164 -I--A--R--Y--I--R--L--H--P--T--H--Y--S--I--R--S--T--L--R--M--

Exon 24

6721 GAGTTGATGGGCTGTGATTTAAATA^SGTTCAGCATGCCATTGGGAATGGA^YCAGTAAAGCA
6550 GAGTTGATGGGCTGTGATTTAAATA^SGTTCAGCATGCCATTGGGAATGGAGAGTAAAGCA
2184 -E--L--M--G--C--D--L--N--S--C--S--M--P--L--G--M--E--S--K--A--

6781 ATATCAGATGCACAGATTACT^SGCTTCATCCTACTTTACCAATATGTTTGCCACCTGGTCT
6610 ATATCAGATGCACAGATTACTGCTTCATCCTACTTTACCAATATGTTTGCCACCTGGTCT
2204 -I--S--D--A--Q--I--T--A--S--S--Y--F--T--N--M--F--A--T--W--S--

Exon 25

6841 CCTTCAAAGCT^{SD}CGACTTCACCTCCAAGGGAGGAGTAATGCCTGGAGACCTCAGGTGAAT
 6670 CCTTCAAAGCTCGACTTCACCTCCAAGGGAGGAGTAATGCCTGGAGACCTCAGGTGAAT
 2224 -P--S--K--A--R--L--H--L--Q--G--R--S--N--A--W--R--P--Q--V--N--

6901 AATCCAAAAGAGTGG^KCTGCAAGTGGACTTCCAGAAGACA^RATGAAAGTCACAGGAGTAACT
 6730 AATCCAAAAGAGTGGCTGCAAGTGGACTTCCAGAAGACAATGAAAGTCACAGGAGTAACT
 2244 -N--P--K--E--W--L--Q--V--D--F--Q--K--T--M--K--V--T--G--V--T--

6961 ACTCA^RGGGAGTAAAACTCTCTGCTTACCAGCATGTATGTGAAGGAGTTCCTCATCTCCAGC
 6790 ACTCAGGGAGTAAAACTCTCTGCTTACCAGCATGTATGTGAAGGAGTTCCTCATCTCCAGC
 2264 -T--Q--G--V--K--S--L--L--T--S--M--Y--V--K--E--F--L--I--S--S--

Exon 26

7021 AGTCAAGATGGCCATCAGTGGACTCTCTTTTTTTCAGAATGGCAAAGTAAAGGTTTTTTCAG
 6850 AGTCAAGATGGCCATCAGTGGACTCTCTTTTTTTCAGAATGGCAAAGTAAAGGTTTTTTCAG
 2284 -S--Q--D--G--H--Q--W--T--L--F--F--Q--N--G--K--V--K--V--F--Q--

7081 GGAAATCAAGACTCCTTCACACCTGTGGTGAACCTCTCTAGACCCACCGTTACTGACTCGC
 6910 GGAAATCAAGACTCCTTCACACCTGTGGTGAACCTCTCTAGACCCACCGTTACTGACTCGC
 2304 -G--N--Q--D--S--F--T--P--V--V--N--S--L--D--P--P--L--L--T--R--

7141 TACCTTCGAATTCACCCCCAGAGTTGGGTGCACCAGATTGCCCTGAGGATGGAGGTTCTG
 6970 TACCTTCGAATTCACCCCCAGAGTTGGGTGCACCAGATTGCCCTGAGGATGGAGGTTCTG
 2324 -Y--L--R--I--H--P--Q--S--W--V--H--Q--I--A--L--R--M--E--V--L--

7201 GGCTGCGAGGCACAGGACCTCTACTGAGGG
 7030 GGCTGCGAGGCACAGGACCTCTACTGA
 2344 -G--C--E--A--Q--D--L--Y--*-

Y

7231 TGGCCACTGCAGCACCTGCCACTGCCGTCACCTCTCCCTCCTCAGCTCCAGGGCAGTGTC

7321 CCTCCCTGGCTTGCCTTCTACCTTTGTGCTAAATCCTAGCAGACACTGCCTTGAAGCCTC

7381 CTGAATTAACTATCATCAGTCCTGCATTTCTTTGGTGGGGGGCCAGGAGGGTGCATCCAA

7441 TTTAACTTAACTCTTACCTATTTTCTGCAGCTGCTCCCAGATTACTCCTTCCTTCCAATA

26BF

7501 TAACTAGGCAAAAAGAAGTGAGGAGAAACCTGCATGAAAGCATTCTTCCCTGAAAAGTTA

26AR

7561 GGCCTCTCAGAGTCACCACTTCCTCTGTTGTAGAAAAACTATGTGATGAACTTTGAAAA

7621 AGATATTTATGATGTTAACATTTTCAGGTAAAGCCTCATACGTTTAAAATAAACTCTCAG

7681 TTGTTTATTATCCTGATCAAGCATGGAACAAAGCATGTTTCAGGATCAGATCAATACAAT

7741 CTTGGAGTCAAAAGGCAAATCATTTGGACAATCTGCAAAATGGAGAGAATACAATAACTA

7801 CTACAGTAAAGTCTGTTTCTGCTTCCTTACACATAGATATAATTATGTTATTTAGTCATT

7861 ATGAGGGGCACATTCTTATCTCCAAACTAGCATTCTTAAACTGAGAATTATAGATGGGG

R *

7921 TTCAAGAATCCCTAAGTCCCCTGAAATTATATAAGGCATTCTGTATAAATGCAAATGTGC

7981 ATTTTTCTGACGAGTGTCATAGATATAAAGCCATTGAGTCTTAATTCTGACCAATAAAA

8041 AAATAAGTCAGGAGGATGCAATTGTTGAAAGCTTTGAAATAAAAATAACAATGTCTTCTTG

26BR

8101 AAATTTGTGATGGCCAAGAAGAAAAATGATGATGACATTAGGCTTCTAAAGGACATACAT

8161 TTAATATTTCTGTGGAAATATGAGGAAAATCCATGGTTATCTGAGATAGGAGATACAAAC

8221 TTTGTAATTCTAATAATGCACTCAGTTTACTCTCTCCCTCTACTAATTCCTGCTGAAAA

8281 TAACACAACAAAAATGTAACAGGGGAAATTATATACCGTGACTGAAAAGTAGAGTCCTAC

 **

8341 TTACATAGTTGAAATATCAAGGAGGTCAGAAGAAAATTGGACTGGTGAAAACAGAAAAAA

8401 CACTCCAGTCTGCCATATCACCACACAATAGGATCCCCCTTCCTTGCCCTCCACCCCCATA

 Y 26CF

8461 AGATTGTGAACGGTTTACTGCTCCTTCCATCTGCCTGACCCCTTCACTATGACTACACAG

 26DR R

8521 AATCTCCTGATAGTAAAGGGGGCTGGAGACAAGGATAAGTTATAGAGCAGTTGGAGGAAG

8581 CATCCAAAGATTGCAACCCAGGGCAAATGGAAAACAGGAGATCCTAATATGAAAGAAAA

 R

8641 TGGATCCCAATCTGAGAAAAGGCAAAAAGAATGGCTACTTTTTTCTATGCTGGAGTATTTT

8701 CTAATAATCCTGCTTGACCCTTATCTGACCTCTTTGGAAACTATAACATAGCTGTCACAG

 Y

8761 TATAGTCACAATCCACAAATGATG^MCAGGTGCAAATGGTTTATAGCCCTGTGAAGTTCTTA
.....
.....

8821 AAGTTTAGAGGCTAACTTACAGAAATGAATAAGTTGTTTTGTTTTATAGCCCGGTAGAGG
.....
.....

8881 AGTTAA^{26EF}CCCCAAAGGTGATATGGTTTTATTTCCCTGTTATGTTTAACTT^RGATA^{26DR}AATCTTATT
.....
.....

8941 TTGGCATTCTTTTCCCATTGACTATATACATCTCTATTTCTCAAATGTTTCATGGAAGTAC
.....
.....

9001 CTCTTTTATTTTCCTGCTGGTTTCTTCAGTAATGAGTTAAATAAAACATTGACACATAC
.....
.....

26ER

Aaacaatgcctttgagaattgtggtttttacactggaataaaaaatgtgaacactgattt
ttaaaacaaataggggcactgaatagcaagatggacactctagaaaaccaaattagttag
ttagaaaaccagattaaattgaactcagagtaaaaaatgataaattcatgagagctgaa
taaaataaatcagaaatggagcctcaatccaggagaacagcttatatggagagagagaga
ctgagagagaaaatgggagtttctgttcaatgggcataaagtttcagctatgctgctgggc
acagtggctcatgcctgtaatcctggcactttgggaggccaaggcggcggtacacctga
ggtcgggagttcaagaccagcctagccaacatggcgaaatcccgtcgtataaaaaatat
aaaaattagttgggcatggtggcacatgcctgtagtcccagctacttgggaggttgaggc
acaagaatcgcttgaaccgggagggcgggtggaggttcagtgagccaagatcggtgccgt
tgcattccatcagcctgggcaacagagtgagactccatctcaaaaaaaaaaagtttcag
ttatgcaagatgaattaagttctacagatctaattgtacaacattgtgcatatagctaacg
ataccatactgggcacttaaaaatttggttaagagagcagataccacggttggttcttact
acagttttacaacaatgcaaaaatgcattgttaacactatgtgcaaatagaatagcaaag
aatgtatacagtgcccatgagcataacattctagtggaccacaatgcatcagagttcag
tgagactcaaagggagaagaatgttggcaaagtcagtggggaggagaaaatgagtgatgat
gaaataggcgggtctcattaaacttcccttacaaaacacaagatcaaagataaaattatta
agaatttcgagacagtgacatcagagcaaacgaaacatag